

## University of Groningen

### Local properties in modal logic

van Ditmarsch, Hans; van der Hoek, Wiebe; Kooi, Barteld

*Published in:*  
Artificial Intelligence

*DOI:*  
[10.1016/j.artint.2012.04.007](https://doi.org/10.1016/j.artint.2012.04.007)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2012

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
van Ditmarsch, H., van der Hoek, W., & Kooi, B. (2012). Local properties in modal logic. *Artificial Intelligence*, 187, 133-155. <https://doi.org/10.1016/j.artint.2012.04.007>

#### Copyright

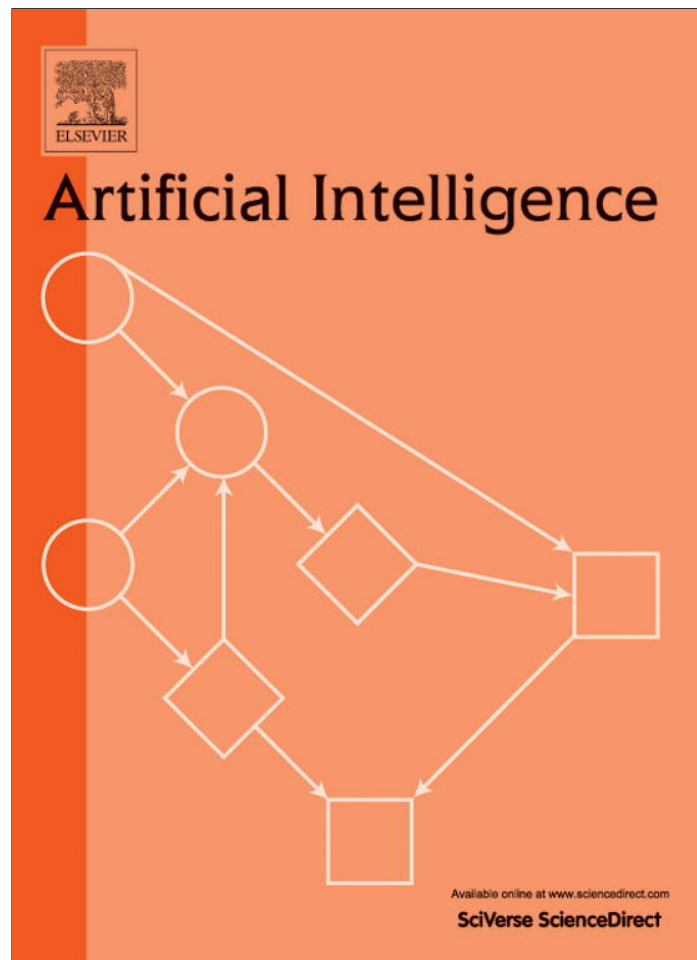
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

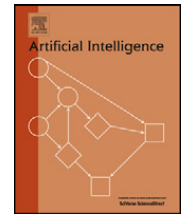
<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

## Artificial Intelligence

www.elsevier.com/locate/artint



## Local properties in modal logic

Hans van Ditmarsch<sup>a</sup>, Wiebe van der Hoek<sup>b,\*</sup>, Barteld Kooi<sup>c</sup><sup>a</sup> Department of Logic, University of Seville, Spain<sup>b</sup> Department of Computer Science, University of Liverpool, UK<sup>c</sup> Faculty of Philosophy, University of Groningen, The Netherlands

## ARTICLE INFO

## Article history:

Received 18 August 2011

Received in revised form 1 March 2012

Accepted 22 April 2012

Available online 2 May 2012

## Keywords:

Knowledge representation

Modal logic

Correspondence

Canonicity

Local properties

Epistemic logic

## ABSTRACT

In modal logic, when adding a syntactic property to an axiomatisation, this property will semantically become true in all models, in all situations, under all circumstances. For instance, adding a property like  $K_a p \rightarrow K_b p$  (agent  $b$  knows at least what agent  $a$  knows) to an axiomatisation of some epistemic logic has as an effect that such a property becomes *globally* true, i.e., it will hold in all states, at all time points (in a temporal setting), after every action (in a dynamic setting) and after any communication (in an update setting), and every agent will know that it holds, it will even be common knowledge. We propose a way to express that a property like the above only needs to hold *locally*: it may hold in the actual state, but not in all states, and not all agents may know that it holds. We achieve this by adding relational atoms to the language that represent (implicitly) quantification over all formulas, as in  $\forall p(K_a p \rightarrow K_b p)$ . We show how this can be done for a rich class of modal logics and a variety of syntactic properties. We then study the epistemic logic enriched with the syntactic property 'knowing at least as much as' in more detail. We show that the enriched language is not preserved under bisimulations. We also demonstrate that adding public announcements to this enriched epistemic logic makes it more expressive, which is for instance not true for the 'standard' epistemic logic **S5**.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Modal logic has become *the* framework for formalising areas in computer science and artificial intelligence as diverse as distributed computing [14], reasoning about programs [15], verifying temporal properties of systems [17], game theoretic reasoning [25], and specifying and verifying multi-agent systems [31]. Regarding the latter example alone, since Moore's pioneering work [19] on knowledge and action, agent theories like intention logic [5] and BDI [20] use modal logic (where the modalities represent time, action, informational attitudes like knowledge or belief, or motivational attitudes like desires or intentions) to analyse interactions between modalities, like *perfect recall*, *no-learning*, *realism*, or different notions of *commitment*. As for epistemic modal logic, since the seminal work of Hintikka [16], modal epistemic logic has played a key role in knowledge representation, witnessed by the literature on reasoning about knowledge in computer science [7], and artificial intelligence [18]. The current activities in dynamic epistemic logic [1,27] can be seen as providing a modal logical analysis in the area of belief revision, thereby providing it with a natural basis for multi-agent belief revision, giving an account of the change of higher order information, and capturing this all in one and the same object language: a modal language, indeed.

The popularity of modal logic in those areas is partly explained by its appealing *semantics*: the notion of state is a very powerful one when it comes to modeling computations of a machine, or describing possibilities that an agent

\* Corresponding author.

E-mail addresses: hvd@us.es (H. van Ditmarsch), wiebe@csc.liv.ac.uk (W. van der Hoek), B.P.Kooi@rug.nl (B. Kooi).

thinks/desires/fears to be possible. Another strong feature of modal logic is its *flexibility*: the fact that temporal, dynamic, informational and motivational attitudes can be represented by modalities does not mean that they all satisfy the same laws. Rather, depending on the interpretation one has in mind, one can decide to either embrace or abandon certain principles for each of the modalities used. Syntactically, this means one assumes a number of axioms or inference rules for a modality or for the interaction of some modalities, and more often than not, this semantically corresponds to assuming some specific properties of the associated accessibility relations in the corresponding models.

In the context of epistemic logic for instance, adding specific modal axioms allows one to specify that the knowing agent is *veridical* ( $K_a p \rightarrow p$ ): if agent  $a$  knows that  $p$ , then  $p$  must be true, or that he is positively ( $K_a p \rightarrow K_a K_a p$ ) or negatively ( $\neg K_a p \rightarrow K_a \neg K_a p$ ) *introspective*. Those axioms happen to correspond (in a precise way: correspondence theory for modal logic is already some decades old, cf. [23]) to reflexivity, transitivity and Euclidicity of the associated accessibility relation  $R_a$ , respectively. Moreover, the axioms are *canonical* for it: adding the syntactic axiom to a modal logic enforces the canonical model for the logic to have the corresponding property, which then in turn implies that completeness of the logic with respect to the class of models satisfying that relational property is guaranteed. At this point, it is important to note the difference between  $K_a p \rightarrow p$  as a *formula* and that as a *scheme*, or *axiom*: as a formula, it merely expresses that regarding the atom  $p$ , agent  $a$  does not know it without it being true. However, when we assume it as an axiom, or as a scheme, it means that we declare it to hold for every substitution instance of  $p$ , in other words, we assume that for all formulas  $\varphi$ , the implication  $K_a \varphi \rightarrow \varphi$  holds.

It is often argued (indeed, already by Hintikka in [16]) that a distinguishing feature between knowledge and belief is that whereas knowledge is veridical, belief need not be, i.e., the scheme  $B_a p \rightarrow p$  should not be assumed as an axiom for belief. This then simply entails that epistemic logics have veridicality as an axiom, and doxastic logics have not. Semantically speaking: the accessibility relations denoting knowledge are reflexive, those denoting belief need not be. But how then to deal with a situation where we want to express that “currently,  $a$ ’s beliefs happen to be true”? If we add  $B_a p \rightarrow p$  as an axiom to our logic, the effect is that in all models (with respect to which the logic is complete), and in all states, all instances of that axiom are true, i.e., for all models  $M$ , for all states  $s$  and for all formulas  $\varphi$ , we then have  $M, s \models B_a \varphi \rightarrow \varphi$ . Given a model  $M$  and a state  $s$  we can express that  $a$ ’s belief that an individual proposition  $q$  holds is correct:  $M, s \models B_a q \wedge q$ . And we can express that  $a$ ’s belief about  $q$  is correct:  $M, s \models (B_a q \rightarrow q) \wedge (B_a \neg q \rightarrow \neg q)$ . But what we *cannot* express in modal logic is that  $B_a \varphi \rightarrow \varphi$  holds for all  $\varphi$  in one state, without claiming at the same time it should hold throughout the model. As a consequence, we cannot express in the object language that agent  $b$  thinks that agent  $a$ ’s beliefs are correct, while agent  $c$  believes that  $a$  is wrong about a proposition  $q$ . The closest one gets to expressing that would be to say that for all  $\varphi$ , in  $M, s$  we have  $M, s \models B_b (B_a \varphi \rightarrow \varphi) \wedge B_c ((B_a q \wedge \neg q) \vee (B_a \neg q \wedge q))$  (but here, the quantification over  $\varphi$  is on a meta-level, and not in the scope of  $B_b$ ). Neither can we say, in a temporal doxastic context, that  $a$ ’s beliefs now are correct, but tomorrow they need not be.

To give another example of the same phenomenon, suppose one adds the scheme  $K_a p \rightarrow K_b p$  to a modal logic ( $b$  knows everything that  $a$  knows). Semantically, this means  $R_b \subseteq R_a$ . If the logic is about a set of agents  $A$ , then it becomes common knowledge among  $A$  that  $b$  knows at least what  $a$  knows! And if there is a notion of time, we have that it will always be the case that  $b$  knows at least what  $a$  knows, and, when having modalities for actions, it follows that no action can make it come about that  $a$  has a secret for  $b$ , in particular, it is impossible to inform  $a$  about something that  $b$  does not already know—this rules out dynamics which are, in contrast, very possible in dynamic epistemic logic.

So, the general picture in modal logic that we take as our starting point is the following. One has a modal logic to which one adds an axiom scheme  $\theta$  (say,  $B_a p \rightarrow p$ ). If one is lucky, the scheme corresponds to a relational property  $\Theta(x)$  (in the case above,  $R_{xx}$ ). However, adding  $\theta$  to the logic means having  $\Theta(x)$  true everywhere, implying that  $\theta$  is always true. What we are after is looking at ways to enforce the scheme  $\theta$  *locally*. To do so, we will add a marker  $\Box$  to the modal language, such that  $\Box$  is true locally, in a state  $s$ , if and only if  $\Theta$  is true, locally (i.e.,  $R_{ss}$  holds).

In [28], in the context of a multi-agent logic **S5**, this is done for the scheme ‘knowing at least as much as’. The expression  $a \succcurlyeq b$  in [28], when true at  $w$  means formally ‘ $a$  considers at least as many accessible worlds from  $w$  as  $b$ ’, and informally ‘ $a$  is at least as uncertain as  $b$  about the actual state of affairs at  $w$ ’, is an example of such a marker  $\Box(a, b)$ , named *Sup*( $a, b$ ) here, and in this case  $\Theta(a, b)(x)$  is the property  $\forall y (R_b xy \Rightarrow R_a xy)$ . The results of [28] are generalised in [29] to more general modal logics **K**( $+\varphi_1, \dots, +\varphi_n$ ) for formulas  $\varphi_i$  satisfying some additional condition, and this is also the main focus of our current contribution.

It is also possible to add several markers at the same time. This then enables that not only can we make global properties locally true, but it also allows for more subtle quantifications over formulas than is allowed in modal logic. This makes it possible to express properties like “If all of John’s beliefs are correct, then so must Mary’s beliefs be”, or “If John knows now everything that Mary knows, then that must have been true yesterday as well” or “If John’s beliefs are correct, then he must know that Mary’s beliefs are correct as well”. The quantification needed for the latter —  $\forall \varphi (B_j \varphi \rightarrow \varphi) \rightarrow K_j \forall \psi (B_m \psi \rightarrow \psi)$  — cannot even be achieved by adding an axiom! For more examples of such quantification, see Section 2.1.

This paper is organised as follows. In Section 1.1 we outline our approach towards a modal logic with local schemes and we explain how, for a variety of property schemes  $\theta$ , one can enforce them to be locally true. To understand our approach, one needs to carefully distinguish between three formal languages, which we formally define in Section 2. Then, in Section 3, we provide a general way to enrich a modal logic with relational atoms, in such a way that the resulting logic is sound and complete with respect to models where those atoms define, in a local sense, to a first order property. In Section 4 we then zoom in on the specific relational atom  $a \succcurlyeq b$ , which locally specifies that  $R_a \supseteq R_b$ . In particular, we give the modalities in

this section an epistemic interpretation, so that  $a \succcurlyeq b$  captures  $K_a p \rightarrow K_b p$  as a local property, i.e., we study such a scheme locally within a modal epistemic logic. Finally, in Section 5 we summarise and conclude.

### 1.1. Towards a modal logic with local schemes

In this paper, we describe two languages to reason about Kripke models. The place where these languages meet is important for our set-up. Let us outline the overall approach with the aim of an example: formal definitions follow in Section 2. First of all, we are interested in a *modal scheme*  $\theta(a, b, p)$ , which in our example reads  $[a]p \rightarrow [b]p$  in a modal language  $\mathcal{L}$  (generally, we write  $[a]\varphi$  for modal formulas, but for epistemic interpretations we may write  $K_a\varphi$ , and for doxastic ones  $B_a\varphi$ ). To the modal language we add a relational atom  $\Box(a, b)$ , or, in this specific case  $Sup(a, b)$ , which will be true in a state  $s$  iff  $\forall y(R_bsy \Rightarrow R_asy)$  holds. The latter property is a formula  $\Theta(a, b)(s)$  in a first-order language  $\mathcal{L}^1$ .

Our modal logic should now formalise the idea that  $\theta(a, b, p)$  and  $\Box(a, b)$  ‘capture the same’. Indeed, we will add  $\Box(a, b) \rightarrow \theta(a, b, p)$  (in our example:  $Sup(a, b) \rightarrow ([a]p \rightarrow [b]p)$ ) as an axiom to the logic. However, we will argue that one cannot just add the opposite direction as an axiom. Instead, the logic will take on an inference rule that ensures that something along the following lines holds: consistency of a formula  $\varphi$  with an occurrence of  $\neg \Box(a, b)$  is the same as consistency of  $\varphi$  with the occurrence of  $\neg \Box(a, b)$  replaced by  $\neg \theta(a, b, p)$  (where  $p$  is a fresh atom). For completeness of the logic, we then take care that in its canonical model, the truth of  $\theta(a, b, p)$  in a specific world (i.e., maximal consistent set  $\Delta$ ) coincides with property  $\Theta(\Delta)$ .

The languages that we define are simple extensions of languages usually studied in standard modal logic [3,4]. More specifically, our modal logic extends that of modal logic with some relational atoms  $\Box$ , and the first order language is the standard language to reason about properties of accessibility relations. Our completeness proof, in turn, is an extension of ‘standard’ completeness proofs in modal logic: we sometimes have to add fresh atoms  $p$  to ensure that  $\theta(a, b, p)$  is satisfied. We have borrowed ideas from [6] to prove our Extension Lemma 2 and ideas from [9–11,21] to make this lemma work ‘everywhere in the canonical model’. Finally, from [11,12] we borrow the notion of  $r$ -persistence (imposed on  $\theta_{\Box}(\vec{a}, \vec{p})$  here) to make our completeness proof work.

## 2. Language and semantics

As outlined above, we deal with two languages, which are all interpreted over the same objects, i.e., Kripke models. The languages are an extended modal language  $\mathcal{L}$ , and a first order language  $\mathcal{L}^1$ . For both languages, we assume a (finite, although this is not crucial for our results) set of modality labels  $A = \{a_1, \dots, a_{|A|}\}$ . In the modal language, these will give rise to modalities  $[a]$ , and in the other language, we assume to have a binary relation  $R_a$  for each  $a \in A$ . For the latter language, we also assume to have a set of variables  $\mathcal{X} = \{x, y, \dots\}$ . The variables will range over possible worlds: note that in  $\mathcal{L}^1$  we do not assume to have constants. For the modal language  $\mathcal{L}$  we assume a finite set  $\rho = \{\Box_1, \Box_2, \dots, \Box_m\}$  of relational atoms: they are nothing else than atomic symbols of which the truth depends on local properties of accessibility relations (see the function  $I$  in Definition 3). Therefore, we will often write  $\Box(a_1, \dots, a_n)$  rather than  $\Box$  to make this dependence clear, and treat  $\Box$  as if it were an  $n$ -ary relational *predicate* (rather than an *atomic* symbol). Our languages will be denoted  $\mathcal{L}(A, \pi, \rho)$  (the modal language), and  $\mathcal{L}^1(A, \mathcal{X})$  (the first order language). If the parameters for the languages are clear, we will also write  $\mathcal{L}$ , and  $\mathcal{L}^1$ , respectively.

**Definition 1** (Modal language). Let the sets  $A, \pi$ , and  $\rho$  be as described above. The modal language  $\mathcal{L}(A, \pi, \rho)$  is defined as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [a]\varphi \mid \Box(a_1, \dots, a_n)$$

where  $a, a_1, \dots, a_n \in A$ ,  $p \in \pi$  and  $\Box$  is an  $n$ -ary relational atom in  $\rho$ . Formula  $\langle a \rangle\varphi$  is shorthand for  $\neg[a]\neg\varphi$  and we also assume the usual definitions for disjunction, implication and bi-implication. If the modality is an epistemic one, the labels denote agents, and we write  $K_a\varphi$  rather than  $[a]\varphi$ . For a doxastic interpretation we write  $B_a\varphi$ , etc.

A formula without occurrences of relational atoms is called a purely modal formula. Suppose we have a multi-modal formula  $\theta(a_1, \dots, a_n, p_1, \dots, p_k)$  where  $a_1, \dots, a_n$  are labels of modalities  $[a_1], \dots, [a_n]$  and  $p_1, \dots, p_k$  are atoms. We will write  $\vec{a}$  for the tuple  $a_1, \dots, a_n$  and  $\vec{p}$  for  $p_1, \dots, p_k$ . When we write  $a \in \vec{a}$  we mean that  $a$  is one of the labels occurring in the tuple  $\vec{a}$ , likewise for  $p$  and  $\vec{p}$ . Finally, for any tuple  $\vec{x} = x_1, \dots, x_n$  with each  $x_i$  taken from some set  $X$ , we will write  $\vec{x} \in \vec{X}$ .

**Definition 2** (First order language). Let  $A$  and  $\mathcal{X}$  be given. First define a language  $\mathcal{L}^+(A, \mathcal{X})$ :

$$\Theta ::= R_a xy \mid \forall y \Theta \mid \neg \Theta \mid \Theta \ \& \ \Theta$$

with  $a \in A$ , and  $x, y \in \mathcal{X}$ . Now, our first order language  $\mathcal{L}^1(A, \mathcal{X})$  is the one-free-variable sublanguage of  $\mathcal{L}^+$ , i.e., the sublanguage of  $\mathcal{L}^+$  consisting of all formulas with one variable not in the scope of a quantifier. If  $\Theta \in \mathcal{L}^1(A, \mathcal{X})$  has  $x$  as its only free variable, and if  $a_1, \dots, a_n$  are all the modality labels occurring in  $\Theta$ , we will also write  $\Theta(\vec{a})(x)$  for  $\Theta$ .

As mentioned earlier, both languages will be interpreted over Kripke models.

**Table 1**

In this table,  $\vec{a}$  is a sequence  $a$  or  $(a, b)$  or  $(a, b, c)$  of modality labels, and  $\vec{p}$  is either the single atom  $p$  or the sequence  $p, q$ .  $\Theta(\vec{a})(x)$  is a property of a state  $x$ , and  $\Box(\vec{a})$  is a name in the object language such that  $\Box(\vec{a})$  holds at  $w$  iff  $\Theta(\vec{a})(w)$  holds of  $M$ .

$\Theta(\vec{a}, \vec{p})$	$\Theta(\vec{a})(x)$	$\Box(\vec{a})$
$[a]p \rightarrow [b]p$	$\forall y(R_bxy \Rightarrow R_axy)$	$Sup(a, b)$
$[c]p \rightarrow [a][b]p$	$\forall y, z((R_axy \ \& \ R_byz) \Rightarrow R_cxz)$	$Trans(a, b, c)$
$\neg[a]\bot$	$\exists y R_axy$	$Ser(a)$
$[a]p \rightarrow p$	$R_axx$	$Refl(a)$
$\langle a \rangle p \rightarrow [b]\langle c \rangle p$	$\forall yz((R_axy \ \& \ R_bxz) \Rightarrow R_cyz)$	$Eucl(a, b, c)$
$\langle a \rangle p \rightarrow \langle b \rangle \langle c \rangle p$	$\forall z(R_axz \Rightarrow \exists y R_bxy \ \& \ R_czy)$	$Dens(a, b, c)$
$\langle \langle a \rangle p \ \& \ \langle b \rangle q \rangle \rightarrow \langle c \rangle (p \ \& \ q)$	$\forall y, z((R_axy \ \& \ R_bxz) \Rightarrow (y = z \ \& \ R_cxy))$	$Func(a, b, c)$

**Definition 3** (*Kripke models and frames*). Given  $A, \pi$  and  $\rho$ , a Kripke model for  $A, \pi$  and  $\rho$  (or, a Kripke model, if  $A, \pi$  and  $\rho$  are clear, or not relevant) is a tuple  $M = \langle W, R, I, V \rangle$  where

- $W$  is a set of possible worlds, also called states
- $R : A \rightarrow \wp(W \times W)$  assigns a binary relation to each modality label
- $I : \rho \rightarrow \mathcal{L}^1(A, \mathcal{X})$  assigns a first order property to each relational atom in  $\rho$
- $V : \pi \rightarrow \wp(W)$  assigns a set of possible worlds to each propositional variable

Rather than  $(w, v) \in R(a)$  we will write  $R_a wv$ . For  $M = \langle W, R, I, V \rangle$  and  $w \in W$ , we let  $R_a(w)$  denote  $\{v \in W \mid R_a wv\}$ . A Kripke frame is a tuple  $F = \langle W, R, I \rangle$  such that  $\langle M, V \rangle = \langle W, R, I, V \rangle$  is a model. The ‘arity’ of a symbol  $\Box \in \rho$  can be read off from its interpretation  $I(\Box)$ : if  $I(\Box)$  refers to modalities  $a_1, \dots, a_n$ , then we may write  $\Box(\vec{a})$  for  $\Box$ . A pointed Kripke model is a pair  $(M, w)$  where  $w \in W$  is a designated point (‘the actual world’); we often delete the parentheses around it. We will sometimes use the terminology of  $I$ -models and  $I$ -frames to highlight the additional interpretation  $I$  compared to ‘standard’ [3] Kripke models.

**Definition 4** (*Semantics of modal formulas*). Let  $A$  and  $\pi$  be given. Also, let  $M = \langle W, R, I, V \rangle$ . Then we define, for  $\varphi \in \mathcal{L}(A, \pi, \rho)$ :

$$\begin{aligned}
 M, w \models p & \quad \text{iff} \quad w \in V(p) \\
 M, w \models \neg\varphi & \quad \text{iff} \quad M, w \not\models \varphi \\
 M, w \models \varphi \wedge \psi & \quad \text{iff} \quad M, w \models \varphi \text{ and } M, w \models \psi \\
 M, w \models [a]\varphi & \quad \text{iff} \quad \text{for all } v \text{ if } R_a wv, \text{ then } M, v \models \varphi \\
 M, w \models \Box(\vec{a}) & \quad \text{iff} \quad I(\Box(\vec{a}))(w) \text{ holds}
 \end{aligned}$$

The class of all models over  $A, \pi$  and  $\rho$  is denoted  $\mathcal{K}(A, \pi, \rho)$ . All models with interpretation  $I$  are denoted  $\mathcal{K}(A, \pi, \rho, I)$ . Validity of  $\varphi$  in a model  $M$  (denoted by  $M \models \varphi$ ) is defined as usual. Moreover,  $\mathcal{K}(A, \pi, \rho) \models \varphi$  means that for all models  $M = \langle W, R, I, V \rangle$  over  $A, \pi$  and  $\rho$ , we have  $M \models \varphi$ . If  $F = \langle W, R, I \rangle$  is a frame,  $F, w \models \varphi$  is defined as: for all valuations  $V, \langle W, R, I, V \rangle, w \models \varphi$ . To distinguish our models and frames from how they are usually defined in modal logic, we will sometimes refer to them as  $I$ -models and  $I$ -frames, respectively.

Interpretation of  $\mathcal{L}^1(A, \mathcal{X})$ -formulas in a model  $M = \langle W, R, I, V \rangle$  is straightforward. For  $\mathcal{L}^2(A, \Pi, \mathcal{X})$ , we assume that  $P(s)$  holds for a predicate  $P$  iff  $s \in V(p)$ . In other words, the link between a propositional atom and a unary predicate is implicit by using lower-case and upper-case notation.

**Example 1.** We give three examples: more are provided in Table 1.

1. Let  $\Box(a, b)$  be such that in  $M$  with interpretation  $I$ , we have  $I(\Box(a, b)) = \Theta(a, b)$  where  $\Theta(a, b)(x) = \forall y(R_bxy \Rightarrow R_axy)$ , saying that in the current world  $w$ , the set of  $a$ -successors of  $w$  is a superset of the set of  $b$ -successors of  $w$ . If this is the interpretation of  $\Box(a, b)$ , we will also write  $Sup(a, b)$ .
2. As a second example, take  $\Box = \Box(a)$  to be such that  $I(\Box(a))(x) = R_axx$ . Note that  $B_a \Box(a)$  can hence be interpreted as ‘ $a$  believes that his beliefs are correct’, since  $M, w \models B_a \Box(a)$  does entail that for all  $\varphi$ ,  $M, w \models B_a(B_a\varphi \rightarrow \varphi)$  (but see Remark 1).
3. Finally, take  $\Box(a, b, c)$  with  $I(\Box(a, b, c))(x) = \forall y\forall z((R_axy \ \& \ R_byz) \Rightarrow R_cxz)$ . We will write  $Trans(a, b, c)$  for  $\Theta(a, b, c)$ . Of course, a special case of this is  $\Box = \Box(a, a, a)$  saying that currently, at world  $w$ , the relation  $R_a$  is transitive.

**Remark 1.** Take  $\Box(a)$  and  $M$  such that  $I(\Box(a)) = \forall x R_axx$ . Note that although  $M, w \models \Box(a)$  entails that agent  $a$ ’s beliefs are correct, the converse is not true, as the following example shows (see Fig. 1). Let  $M = \langle W, R, I, V \rangle$  be such that  $W = \{w, u\}$ ,

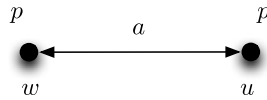


Fig. 1. A model  $M, w$ . States  $w$  and  $u$  verify the same atoms.

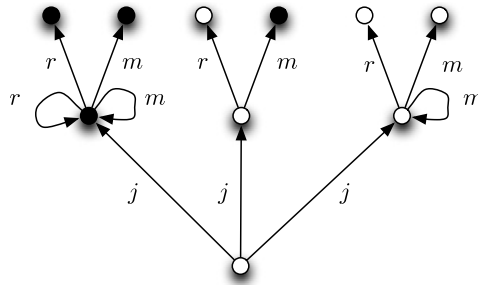


Fig. 2. A model  $M, w$ . The atom  $p$  is true exactly in the worlds that are filled black.

and  $R_a = \{(w, u), (u, w)\}$ . Moreover, assume that for all  $p$ ,  $w \in V(p)$  iff  $u \in V(p)$ . Since  $(M, w)$  and  $(M, u)$  are bisimilar models ([4, Chapters 1 and 5]), we have  $M, w \models \varphi$  iff  $M, u \models \varphi$ , and hence  $M, w \models B_a \varphi \rightarrow \varphi$ , for all purely modal  $\varphi$ . However, since  $(w, w) \notin R_a$ , we have  $M, w \models \neg \Box(\bar{a})$ .

Note that, since  $\Theta(\bar{a})(w)$  does not refer to atomic propositions  $p$  (or, rather predicates  $P$ ), we have that  $\Theta(\bar{a})(w)$  holds in the model  $M = \langle W, R, I, V \rangle$  iff  $\Theta(\bar{a})(w)$  holds in the frame  $F = \langle W, R, I \rangle$ .

#### Definition 5.

1. Let  $\varphi \in \mathcal{L}(A, \pi, \rho)$ , and  $\Phi$  some property which applies to states in a frame.
  - (a) We say that  $\varphi$  defines the frame property  $\Phi$  if for every frame  $F = \langle W, R, I \rangle$  we have  $F \models \varphi$  iff  $\Phi$  holds of  $F$ .
  - (b)  $\varphi$  is said to locally define  $\Phi(x)$  if for every frame  $F$  and every  $w \in W$ , we have  $F, w \models \varphi$  iff  $\Phi(w)$  holds in  $F$ .
2. Let  $\theta_{\Box}(\bar{a}, \bar{p})$  be a purely modal formula,  $\Box(\bar{a}) \in \rho$  and  $\Theta(\bar{a})(x)$  a first order property. If  $\theta_{\Box}(\bar{a}, \bar{p})$  locally defines  $\Theta(\bar{a})(x)$  and  $I$  is such that  $I(\Box(\bar{a})) = \Theta(\bar{a})(x)$ , then we say that  $I$  semantically links  $\theta_{\Box}(\bar{a}, \bar{p})$  and  $\Box(\bar{a})$  through  $\Theta(\bar{a})(x)$ .

Item 1 of Definition 5 generalises a definition of [12, pp. 180, 181] to  $I$ -frames. Our completeness proof relies on even a stronger notion, although there we restrict ourselves to the case of first order properties (i.e., item 2 of Definition 5). The notions defined in the first item are also known as *correspondence* (between a modal formula and some, usually, first order property). There is in fact a rich literature on modal correspondence theory: see for instance the early [23], the chapter on local correspondence in the later [24] and the more recent [4, Chapter 1] and [3, Chapter 3] and the references therein.

Take the specific example in a doxastic context where  $\Theta(a)(x)$  is  $R_a xx$ , and  $I(\Box(a)) = \Theta(a)$ . Note that  $\theta(a, p) = (B_a p \rightarrow p)$  defines  $\Theta(a)(x)$  but still, as shown in Remark 1, the formulas  $\Box(a)$  and  $\theta(a, p)$  are not *equivalent*. Still, the two should be strongly connected, in a sense we will explain in Section 3. We first look at an example, involving our extended modal language.

#### 2.1. A simple example

Consider five friends, Joey, Chandler, Ross, Monica and Phoebe (or  $j, c, r, m$  and  $p$ , for short). In this example, we use ‘think’ and ‘believe’ for the same thing. Joey believes that Monica’s beliefs are at least as accurate as Ross’ beliefs, i.e., Joey believes that if Ross’ beliefs are correct, so must Monica’s be (A). Joey also believes that Monica thinks that Chandler believes anything that Monica believes (B). Although Joey does not think that he believes everything he knows (he thinks that he knows he cannot find a job as an actor, but at the same time cannot believe it), he actually believes anything he knows (C). Moreover, Joey thinks that Chandler’s beliefs are consistent (D). Finally, Joey happens to know that Monica believes that Phoebe is in competition with her for Chandler’s attention, but at the same time Joey thinks that Chandler believes that Phoebe is not in competition with Monica for his attention (E). Then, we conclude that Joey believes that Ross’ beliefs are not guaranteed to be correct (F), or, better, that Joey believes he may assume that some formula is believed by Ross, but not true (F’).

We first give a (semi-formal) formalisation of our assumption using a modal logic that allows for quantification over formulas. Let  $z$  represent the proposition that Joey cannot find himself a job as an actor, and let  $q$  be the proposition that Phoebe is in competition with Monica for Chandler’s attention. This formalisation is given in Table 2, where assumption (A) in our informal description is represented as (a), etc. The formalisation in our language  $\mathcal{L}(A, \pi, \rho)$  follows in Table 3.

We can now be more precise about what it means that our language can do more than just formalising a local version of a global property. For instance, the global property  $B_a p \rightarrow p$  will have a local counterpart  $\text{Refl}(a)$ . Locally, this will denote

**Table 2**

A semi-formal translation of the episode.

(a)	$B_j(\forall\varphi(B_r\varphi \rightarrow \varphi) \rightarrow \forall\varphi(B_m\varphi \rightarrow \varphi))$
(b)	$B_jB_m(\forall\varphi(B_m\varphi \rightarrow B_c\varphi))$
(c)	$\neg B_j(K_jz \rightarrow B_jz) \wedge \forall\varphi(K_j\varphi \rightarrow B_j\varphi)$
(d)	$B_j\forall\varphi(\neg(B_c\varphi \wedge B_c\neg\varphi))$
(e)	$K_jB_mq \wedge B_jB_c\neg q$
(f)	$B_j\neg\forall\varphi(B_r\varphi \rightarrow \varphi)$

**Table 3**

A formalisation of the episode.

(a')	$B_j(\text{Refl}(r) \rightarrow \text{Refl}(m))$
(b')	$B_jB_m\text{Sup}(m, c)$
(c')	$\neg B_j(K_jz \rightarrow B_jz) \wedge \text{Sup}(kj, k)$
(d')	$B_j\text{Ser}(c)$
(e')	$K_jB_mq \wedge B_jB_c\neg q$
(f')	$B_j\neg\text{Refl}(r)$

something that is similar to  $\forall\varphi(B_a\varphi \rightarrow \varphi)$ . But if one looks at our formalisation (a) as given in Table 2, which is our formalisation of the assumption (A) that was given at the beginning of this example, i.e.,  $B_j(\forall\varphi(B_r\varphi \rightarrow \varphi) \rightarrow \forall\varphi(B_m\varphi \rightarrow \varphi))$ , it becomes clear that this is different from the quantification (g) :  $\forall\varphi B_j((B_r\varphi \rightarrow \varphi) \rightarrow (B_m\varphi \rightarrow \varphi))$ , which one would get as a local counterpart of an axiom  $B_j((B_r p \rightarrow p) \rightarrow (B_m p \rightarrow p))$ . That (a) and (g) are not equivalent, can be seen in the model  $M, w$  of Fig. 2, where (a) is true in  $M, w$ , but (g) is not: for the latter,  $\varphi = p$  provides a counter-example. That (a) is true in  $M, w$  is easily seen from realising that  $a$  is formalised by  $a'$  in Table 3.

We then formalise the same episode using the relational atoms  $\Box(\vec{a})$  introduced in Table 1, which results in Table 3. Abusing the language somewhat, we write  $\text{Sup}(kj, j)$  for the relational atom corresponding to  $K_j\varphi \rightarrow B_j\varphi$ —from a language point of view,  $K_j$  and  $B_j$  are simply two different modal operators, say  $[kj]$  and  $[i]$ , respectively.

### 3. Axiomatization

The aim of this section is to provide an axiomatisation for modal logics that are enriched with some relational atoms  $\Box_1(\vec{a}_1), \dots, \Box_m(\vec{a}_m)$ , such that for every  $\Box_k(k \leq m)$ , there is a modal formula  $\theta_{\Box_k}(\vec{a}, \vec{p})$  such that, at least on frames, the two ‘mean the same thing’. In fact, the logic  $\mathbf{K}(A, \pi, \rho, I)$  that we define should be sound and complete with respect to  $\mathcal{K}(A, \pi, \rho, I)$ , so our aim for our logic is that for all formulas  $\varphi \in \mathcal{L}(A, \pi, \rho)$ , the notions  $\mathbf{K}(A, \pi, \rho, I) \vdash \varphi$  and  $\mathcal{K}(A, \pi, \rho, I) \models \varphi$  coincide. The idea to achieve this is as follows. In order to characterise the ‘meaning’ of  $\Box(\vec{a})$ , we first like to specify what follows from it (‘elimination of  $\Box(\vec{a})$ ’): this will be specified by axiom  $\mathbf{Ax}_{\Box}$  in Table 4. Secondly, we need to characterise when one can derive that  $\Box(\vec{a})$  (‘introduction of  $\Box(\vec{a})$ ’), which is our inference rule  $\mathbf{R}_{\Box}$  in that table. To make this all work, we moreover rely on a first order property  $\Theta(\vec{a})$  that can be used as the interpretation of  $\Box(\vec{a})$ , and which is locally defined by  $\theta_{\Box}(\vec{a}, \vec{p})$ . It will turn out that finding such a property is not easy or indeed possible for all  $\theta_{\Box}(\vec{a}, \vec{p})$ , so we will need to impose an additional condition on it: this will be done in Definition 11. This condition may look rather ad-hoc, but, as we will mention in Section 3.3, it is implied by a well-known property of *local r-persistence*. We will provide a formal soundness and completeness result for our logic, and will discuss connections with related approaches and techniques in Section 3.3.

Before plunging in the technical details, in order to get a feel for our axiomatisation, it pays off to recall that a formula of the form  $\Box(\vec{a})$  is a label (like  $\text{Sup}(a, b)$ ), which represents both a modal formula  $\theta_{\Box}(\vec{a}, \vec{p})$  (like  $[a]p \rightarrow [b]p$ ) and a first-order property  $\Theta(\vec{a})(x)$  (like  $\forall y(R_bxy \Rightarrow R_axy)$ ).

This is a good point to remind ourselves of the axiomatisation of the  $\forall$  quantifier in first order logic. Formulated in our context (and quantifying over atomic variables), it has an axiom and an inference rule (let us for simplicity assume that  $p$  coincides with  $\vec{p}$ , i.e., we only have one atom in  $\theta_{\Box}(\vec{a}, \vec{p})$ ):

$$\begin{aligned} \mathbf{Ax}_{\forall} \quad & \forall p \theta_{\Box}(\vec{a}, p) \rightarrow \theta_{\Box}(\vec{a}, p) \\ \mathbf{R}_{\forall} \quad & \text{from } \varphi \rightarrow \theta_{\Box}(\vec{a}, p) \text{ infer } \varphi \rightarrow \forall p \theta_{\Box}(\vec{a}, p) \\ & \text{where } P \text{ does not occur in } \hat{I} \end{aligned}$$

Our axiomatisation is then obtained by (i) choosing  $\Box(\vec{a})$  for  $\theta_{\Box}(\vec{a}, p)$  and (ii), adding a ‘modal component’ to the inference rule, which ensures that  $\theta_{\Box}(\vec{a}, p)$  may be replaced by  $\Box(\vec{a})$  ‘anywhere in the model’.

We now explain our axiom and inference rule in pure modal logic terms, as follows. First of all, suppose that for every relational atom  $\Box(\vec{a})$  and fixed interpretation  $I$  we have a formula  $\theta_{\Box}(\vec{a}, \vec{p})$  such that  $\theta_{\Box}(\vec{a}, \vec{p})$  defines  $I(\Box(\vec{a}))$ . Then, for each  $\Box(\vec{a})$  and related  $\theta_{\Box}(\vec{a}, \vec{p})$  we add an axiom  $\mathbf{Ax}_{\Box}$ , which is  $\Box(\vec{a}) \rightarrow \theta_{\Box}(\vec{a}, \vec{p})$  to our logic  $\mathbf{K}(A, \pi, \rho, I)$ . This makes sense, given  $\mathbf{Ax}_{\forall}$  and our observation above that  $\Box(\vec{a})$  is equivalent to  $\forall P \hat{\Theta}(\vec{a}, P)(x)$ , and the instance  $\hat{\Theta}(\vec{a}, \mathbf{P})(x)$  corresponds to the modal formula  $\theta_{\Box}(\vec{a}, p)$ .



**Table 4**

The axioms and inference rules of the logic  $\mathbf{K}(A, \pi, \rho, I)$ .

<b>Prop</b>	All instances of propositional tautologies
<b>K</b>	$[a](\varphi \rightarrow \psi) \rightarrow ([a]\varphi \rightarrow [a]\psi)$
<b>Ax<math>_{\Box}</math></b>	$\Box(\vec{a}) \rightarrow \theta_{\Box}(\vec{a}, \vec{p})$
<b>MP</b>	From $\varphi \rightarrow \psi$ and $\varphi$ , infer $\psi$
<b>Nec</b>	From $\varphi$ , infer $[a]\varphi$
<b>R<math>_{\Box}</math></b>	From $[s]\theta_{\Box}(\vec{a}, \vec{p})$ infer $[s]\Box(\vec{a})$ where $\vec{p}$ does not occur in $s$
<b>US</b>	From $\varphi$ infer $\varphi[\psi/p]$

Adding the other direction as an implication does not work, as the example  $\Box(\vec{a}) = \Box(a, b) = \text{Sup}(a, b)$  and  $\theta(\vec{a}, \vec{p}) = [a]p \rightarrow [b]p$  shows:  $([a]p \rightarrow [b]p) \rightarrow \text{Sup}(a, b)$  is not a validity: the antecedent may be true due to some specific choice of  $p$ . Note that the modal reading of the rule **R $_{\forall}$**  in our language would give

**R** from  $\varphi \rightarrow \theta_{\Box}(\vec{a}, p)$  infer  $\varphi \rightarrow \Box(\vec{a})$   
where  $p$  does not occur in  $\varphi$

The rule **R** will indeed be derivable in our logic, but we need a little bit more. We will argue how to obtain **R** and its generalisation, **R $_{\Box}$** , by going back to the semantics of our modal language. Suppose that  $\mathcal{K}(A, \pi, \rho, I) \models \varphi \rightarrow ([a]p \rightarrow [b]p)$ , where  $p$  does not occur in  $\varphi$ . This then means that  $\varphi$  must entail that (locally) all  $b$ -successors are  $a$ -successors, i.e.,  $\mathcal{K}(A, \pi, \rho, I) \models \varphi \rightarrow \text{Sup}(a, b)$ , because if the latter would not hold, there would be a model  $M = \langle W, R, I, V \rangle$  such that  $M, w \models \varphi \wedge \neg \text{Sup}(a, b)$ . But since  $p$  does not occur in  $\varphi$ , we could then change the valuation  $V$  for  $p$  freely without changing that of  $\varphi$ , in particular we can choose  $V'$  such that  $x \in V'(p)$  iff  $R_a wx$  (and  $V'(q) = V(q)$  for atoms  $q \neq p$ ). It is easy to see that in the resulting model  $M' = \langle W, R, I, V' \rangle$  we have  $M', w \models \varphi \wedge \neg([a]p \rightarrow [b]p)$ : a contradiction. This means that we need to be able to infer the following in  $\mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I)$ :

$$\begin{aligned} \text{If } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow ([a]p \rightarrow [b]p) \\ \text{then } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow \text{Sup}(a, b), \\ \text{where } p \notin \varphi \end{aligned} \quad (1)$$

The rule (1) is of course a special case of **R $_{\forall}$** . It can also be understood as follows. If  $p$  does not occur in  $\varphi$ , and  $\varphi \rightarrow ([a]p \rightarrow [b]p)$  is true at a state  $s$ , then  $\varphi$  must carry sufficient information such that  $[a]p \rightarrow [b]p$  must hold (it will not be because of specific requirements on  $p$  imposed by  $\varphi$ ) and hence we must have  $\varphi \rightarrow \text{Sup}(a, b)$  holding at  $s$  as well. But in fact we can do the same reasoning that involves successors of  $s$ : suppose  $\varphi$  implies that in all  $R_c$ -successors  $t$  of  $s$ , we have  $M, t \models [a]p \rightarrow [b]p$ . Then (in the same way as for  $s$ ), we must have  $M, t \models \varphi \rightarrow \text{Sup}(a, b)$ . In other words, the following should hold for  $\mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I)$ :

$$\begin{aligned} \text{If } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c]([a]p \rightarrow [b]p) \\ \text{then } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c]\text{Sup}(a, b), \\ \text{where } p \notin \varphi \end{aligned} \quad (2)$$

And the same should hold for all  $R_d$ -successors  $u$  of all  $R_c$ -successors  $t$  of  $s$ , etc. To formalise that a property  $\theta_{\Box}(\vec{a}, \vec{p})$  holds after arbitrary sequences  $\varphi_1 \rightarrow [a_1](\varphi_2 \rightarrow \dots [a_{n-1}](\varphi_n \rightarrow \theta_{\Box}(\vec{a}, \vec{p})) \dots)$ , we follow [28] and introduce *pseudo-modalities*: we will then present an inference rule **R $_{\Box}$**  for every  $\Box \in \rho$  to our axiomatisation  $\mathbf{K}(A, \pi, \rho, I)$ .

**Definition 6** (*Pseudo-modalities*). We define the following *pseudo-modalities*, which are (possibly empty) sequences  $s = ()$  or  $s = (s_1, \dots, s_n)$ , where each  $s_i$  is a formula or a modality label. The formula  $\langle s \rangle \varphi$  represents an  $\mathcal{L}(A, \Pi, \rho)$  formula, as follows (we will omit parentheses around non-empty sequences when they occur with a diamond or box):

$$\begin{aligned} \langle () \rangle \varphi &= \varphi \\ \langle \psi, s_2, \dots, s_n \rangle \varphi &= \psi \wedge \langle s_2, \dots, s_n \rangle \varphi \\ \langle a, s_2, \dots, s_n \rangle \varphi &= \langle a \rangle (\langle s_2, \dots, s_n \rangle \varphi) \end{aligned}$$

We also define  $[s]\varphi$  as  $\neg \langle s \rangle \neg \varphi$ . We say that  $\vec{p}$  does not occur in  $s$  (and write  $\vec{p} \notin s$ ) if none of the atoms  $p$  occurring in  $\vec{p}$  does occur in any of the formulas  $s_i$  in  $s$ .

So, for instance  $\langle a, \psi, b \rangle \varphi$  is an abbreviation of  $\langle a \rangle (\psi \wedge \langle b \rangle \varphi)$ , while  $[a, \psi, b]\varphi$  is short for  $[a](\psi \rightarrow [b]\varphi)$ .

**Definition 7** (Proofsystem). Fix  $A, \pi$  and  $\rho$ . Then, Table 4 comprises the axioms and inference rules of the logic  $\mathbf{K}(A, \pi, \rho, I)$ .

In this table **MP** stands for Modus Ponens, **Nec** for Necessitation, and **US** for Uniform Substitution ( $\varphi[\psi/p]$  stands for substitution of  $\psi$  for every occurrence of  $p$  in  $\varphi$ ). If  $\Box(\bar{a})$  and  $\theta_{\Box}(\bar{a}, \bar{p})$  are connected through the axiom **Ax $_{\Box}$**  and inference rule **R $_{\Box}$** , we say they are axiomatically linked (through axiom **Ax $_{\Box}$**  and rule **R $_{\Box}$** ). If there is a derivation of a formula  $\varphi$  from a set of formulas  $\Gamma$  using  $\Gamma$  and the axioms and inference rules from  $\mathbf{K}(A, \pi, \rho, I)$  we write  $\Gamma \vdash_{\mathbf{K}(A, \pi, \rho, I)} \varphi$ , or  $\Gamma \vdash_{\mathbf{K}} \varphi$ , for short.

**Example 2.** Suppose  $\Box$  is  $\text{Sup}(a, b)$ , so that  $\theta_{\Box}(\bar{a}, \bar{p}) = [a]p \rightarrow [b]p$ . Then Axiom **Ax $_{\Box}$**  becomes  $\Box(a, b) \rightarrow ([a]p \rightarrow [b]p)$ . And the inference rule **R $_{\Box}$**  becomes: from  $\langle s \rangle([a]p \wedge \neg[b]p) \rightarrow \varphi$ , infer  $\langle s \rangle \neg \text{Sup}(a, b) \rightarrow \varphi$ , where  $p$  does not occur in  $\varphi$  or  $s$ .

**Lemma 1.** Let  $\bar{p}$  be a sequence of atoms not occurring in  $\varphi$ . Then rule **R** is an instance of **R $_{\Box}$**  obtained with  $s = ()$ , and **R $_{\Box}$**  is equivalent to **R $^1$**  and **R $^2$** :

- R** From  $\neg \theta_{\Box}(\bar{a}, \bar{p}) \rightarrow \varphi$ , infer  $\neg \Box(\bar{a}) \rightarrow \varphi$
- R $^1$**  If  $\langle s \rangle \neg \Box(\bar{a}) \wedge \varphi$  is consistent,  
then so is  $\langle s \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \varphi$
- R $^2$**  From  $\langle s \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \rightarrow \varphi$ , infer  $\langle s \rangle \neg \Box(\bar{a}) \rightarrow \varphi$

**Theorem 1** (Soundness). For all  $\varphi \in \mathcal{L}(A, \pi, \rho)$ , if  $\mathbf{K}(A, \pi, \rho, I) \vdash \varphi$  then  $\mathcal{K}(A, \pi, \rho, I) \models \varphi$ .

**Proof.** We only show that axiom **Ax $_{\Box}$**  and rule **R $^1$**  are valid, the latter demonstrating validity of **R $_{\Box}$** : all the other axioms and rules are standard. Starting with **Ax $_{\Box}$** , take a model  $M = \langle W, R, I, V \rangle$ , where  $I$  is a parameter of the logic, and  $W, R$  and  $V$  are arbitrary. If  $M, w \models \Box(\bar{a})$ , this means that  $I(\Box(\bar{a}))(w)$  holds. Since  $\theta_{\Box}(\bar{a}, \bar{p})$  characterises  $I(\Box(\bar{a}))$ , we get  $M, w \models \theta_{\Box}(\bar{a}, \bar{p})$ .

Next, we will argue that, if  $\bar{p}$  does not occur in  $s$  and  $\varphi$ , then

If  $\langle s \rangle \neg \Box(\bar{a}) \wedge \varphi$  is satisfiable, then so is  $\langle s \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \varphi$

First of all, note that we can assume that  $s$  never needs two successive agents  $a$  and  $b$  (we can separate them with  $\top$ ) and neither does  $s$  need two successive formulas (we can replace them by their conjunction). Moreover we can assume that  $s$  is either empty or else starts with a modality (a possible condition can be added to  $\varphi$ ). If  $s = ()$ , then the assumption says that  $\neg \Box(\bar{a}) \wedge \varphi$  is satisfiable. That is, there is a model  $M = \langle W, R, I, V \rangle$  and a state  $w \in W$  such that  $M, w \models \neg \Box(\bar{a}) \wedge \varphi$ . The first part of this conjunction says that  $\neg I(\Box(\bar{a}))(w)$ . Since  $\theta_{\Box}(\bar{a}, \bar{p})$  locally defines  $I(\Box(\bar{a}))$ , and, since  $\bar{p}$  does not occur in  $\varphi$ , we can change the valuation  $V$  for  $\bar{p}$  without changing the truth of  $\varphi$ . In other words, we can define a new model  $M' = \langle W, R, I, V' \rangle$  where the truth of  $p \in \bar{p}$  in  $M'$  is such that it falsifies  $\theta_{\Box}(\bar{a}, \bar{p})$ , but if keeps the truth of  $\varphi$  unchanged. We have  $M', w \models \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \varphi$ , demonstrating the satisfiability of  $\langle s \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \varphi$ .

Now we use induction and assume that  $s = \langle c, \psi \rangle \cdot t$ , where  $\psi$  is a formula (not containing  $\bar{p}$ ),  $c$  an arbitrary modality,  $t$  a sequence (not containing  $\bar{p}$ ) and  $\cdot$  denotes concatenation of sequences. Suppose  $\langle \langle c, \psi \rangle \cdot t \rangle \neg \Box(\bar{a}) \wedge \varphi$  is satisfiable. This means that at some  $M, w$ , we have  $M, w \models \varphi$  and for some  $v$  with  $R_c w v$ ,  $M, v \models \langle t \rangle \neg (\Box \bar{a}) \wedge \psi$ . Using induction, we get  $R_c w v$  and  $M, v \models \langle t \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \psi$ , (with  $\bar{p} \notin \psi, t$ ). Hence  $M, w \models \langle \langle c, \psi \rangle \cdot t \rangle \neg \theta_{\Box}(\bar{a}, \bar{p}) \wedge \varphi$ .  $\square$

### 3.1. Back to our example

To formalise the derivation of Table 3, let the set of modalities representing the friends be  $A = \{c, j, m, p, r\}$ , let  $\pi = \{q, z\}$  and let  $\rho = \{\text{Refl}(r), \text{Refl}(m), \text{Sup}(c, m), \text{Ser}(c), \text{Sup}(kj, j)\}$  and those atoms are axiomatically linked with their ‘natural’ modal counterparts (see Table 1 and for  $\text{Sup}(kj, j)$  we take  $K_j p \rightarrow B_j p$ ). Let the resulting logic be  $\mathbf{K}(A, \pi, \rho, I)$ .

First of all, from  $(c')$  and **Ax $_{\text{Sup}(kj, j)}$**  we derive  $K_j B_m q \rightarrow B_j B_m q$ . Together with  $(e')$  this gives  $(e'')$ :  $B_j B_m q \wedge B_j B_c \neg q$ . From  $(d')$ , i.e.,  $B_j \text{Ser}(c)$  and **Ax $_{\text{Ser}(c)}$** , we get  $B_j (B_c \neg q \rightarrow \neg B_c q)$ . Combining this with  $(e'')$  gives  $B_j B_m q \wedge B_j \neg B_c q$ , which is equivalent to  $B_j \neg (B_m q \rightarrow B_c q)$  (\*).

From  $(b')$  and **Ax $_{\text{Sup}(m, c)}$**  we derive  $B_j B_m (B_m p \rightarrow B_c p)$ , for any  $p$  (\*\*). Now, take the formula  $\psi = (B_m q \rightarrow B_c q)$ . From (\*) we have  $B_j \neg \psi$ , and from (\*\*) we conclude  $B_j B_m \psi$ . In other words, we found a formula  $\psi$  for which  $B_j \neg (B_m \psi \rightarrow \psi)$ . Now using the contrapositive of axiom **Ax $_{\text{Refl}(m)}$** , we obtain  $B_j \neg \text{Refl}(m)$ , which, together with  $(a')$ , yields our conclusion  $(f')$ .

Now one may wonder whether this also warrants the conclusion  $(f)$ . Note that we *did* find a formula  $\psi$  above for which we derived that Joe believes that Monica believes it wrongly, but this does not imply that Joey also believes that Ross is mistaken regarding  $\psi$ . Also, although we now have  $(f')$ , as should be clear from Remark 1,  $B_j \neg \text{Refl}(r)$  and  $B_j \neg \forall \varphi (B_r \varphi \rightarrow \varphi)$  are not the same thing. However, what we *do* have is the following. Let  $\varphi$  be  $(a') \wedge (b') \wedge (c') \wedge (d') \wedge (e')$ , and let  $s$  be  $B_j$ , then what we have proven now is

$$\mathbf{K}(A, \pi, \rho, I) \vdash \varphi \rightarrow B_j \neg \text{Refl}(r) \quad (3)$$

Which informally reads that, from our assumptions, we can derive that Joey believes that Ross considers the actual world not (doxastically) possible. If we now add  $(g') = \text{Ser}(j)$  (Joey's beliefs are consistent) to our assumptions  $\varphi' = \varphi \wedge (g')$ , we obtain from (3):

$$\mathbf{K}(A, \pi, \rho, I) \vdash \varphi' \rightarrow M_j \neg \text{Refl}(r) \quad (4)$$

which, with rule  $\mathbf{R}^1$ , gives that, for any  $p$  not occurring in  $\varphi'$ ,  $\varphi' \wedge M_j(B_r p \wedge \neg p)$  is consistent, i.e., we can assume that for some  $p$ , Joey considers it possible that Ross is mistaken in his beliefs regarding *that*  $p$ .

It is worth noting how the axiomatisation makes it possible that some relational atoms (and hence some first-order frame properties) only hold in the scope of a modal operator (like in property (a') and (b') for example): the axiom  $\mathbf{Ax}_{\Box}$  and rule  $\mathbf{R}_{\Box}$  do not require that some relational properties hold, they only specify what should be the case if they hold.

### 3.2. Completeness

In this section we show completeness of the axiomatisation, by the standard method of constructing a canonical model. However, to ensure that the points (maximal consistent sets) in this model are closed under the new inference rule (cf. rule  $\mathbf{R}^1$ ), we have to ensure that those sets are *fully witnessed*, as defined below. We also present a formulation of the condition that we need to impose on the formulas  $\theta_{\Box}(\vec{a}, \vec{p})$  to make the completeness proof work: we will later, in Section 3.3 dwell on this condition further.

**Definition 8.** A theory  $\Gamma$  is a set of formulas. For  $\pi$  a set of propositional atoms,  $\Gamma$  is a  $\pi$ -theory if all propositional atoms in  $\Gamma$  are from  $\pi$ . Given a logic  $\mathbf{L}$ , a theory  $\Gamma$  is  $\mathbf{L}$ -consistent if  $\perp$  cannot be derived from  $\Gamma$  using the axioms and inference rules of  $\mathbf{L}$ . A theory  $\Gamma$  is a maximal  $\mathbf{L}$ -consistent  $\pi$ -theory if it is consistent and no  $\pi$ -theory  $\Delta$  is  $\mathbf{L}$ -consistent while at the same time  $\Gamma \subset \Delta$ . For a logic  $\mathbf{K}(A, \pi, \rho, I)$ , a set of formulas  $\Gamma$  is a *witnessed*  $\pi$ -theory if for every  $\Box(\vec{a})$  and every  $\langle s \rangle \neg \Box(\vec{a}) \in \Gamma$ , there are atoms  $\vec{p}$  such that  $\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p}) \in \Gamma$ , where  $\Box(\vec{a})$  and  $\theta_{\Box}(\vec{a}, \vec{p})$  are axiomatically linked. If  $\Gamma$  is not witnessed, then a formula  $\langle s \rangle \neg \Box(\vec{a})$  for which there is no  $\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p}) \in \Gamma$ , is called a *defect* for the theory  $\Gamma$ . Finally,  $\Gamma$  is said to be *fully witnessed*, if it is witnessed and for every formula of the form  $\langle s \rangle \varphi$ , either that formula or its negation is in  $\Gamma$ .

**Lemma 2 (Extension Lemma).** Let  $\Sigma$  be a  $\mathbf{K}(A, \pi, \rho, I)$ -consistent  $\pi$ -theory. Let  $\pi' \supseteq \pi$  be an extension of  $\pi$  by a countable set of propositional variables. Then there is a maximal  $\mathbf{K}(A, \pi', \rho, I)$ -consistent, witnessed  $\pi'$ -theory  $\Sigma'$  extending  $\Sigma$ .

Before we give a proof, we first define some languages.

**Definition 9.** Let the set of agents  $A$ , the set of atoms  $\pi$  and the set of relational atoms  $\rho$  be fixed. Let  $\mathcal{L}(A, \pi, \rho)$  be as in Definition 1. Let  $\pi^0 = \{p_0, p_1, \dots\}$  be a set of fresh atomic variables, i.e.,  $\pi \cap \pi^0 = \emptyset$  and let  $\pi' = \pi \cup \pi^0$ . Let  $\pi_n = \pi \cup \{p_i \mid i \leq n\}$ . Define  $\mathcal{L}_n$  to be  $\mathcal{L}(A, \pi_n, \rho)$ , and let  $\mathcal{L}_{\omega}$  be  $\mathcal{L}(A, \pi', \rho)$ . A theory  $\Delta \subseteq \Sigma$  is called an *approximation* if for some  $n$  it is a consistent  $\pi_n$ -theory. For such a theory, and any number  $k$ , the sequence  $\vec{p} = \langle p_{n+1}, \dots, p_{n+k} \rangle$  is a new sequence  $\vec{p}$  for  $\Delta$  if  $n$  is the least number such that  $\Delta$  is a  $\pi_n$ -theory.

**Proof of Lemma 2.** Assume an enumeration of  $\psi_0, \psi_1, \dots$  of all formulas of the form  $\langle s \rangle \neg \Box(\vec{a})$ , where  $s$  is a pseudo-modality and  $\Box(\vec{a}) \in \rho$ . Define

$$\Delta^+ = \begin{cases} \Delta \cup \{\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p})\} & \text{where } \vec{p} \text{ is a new sequence for } \Delta, \text{ and } \langle s \rangle \neg \Box(\vec{a}) \text{ is the first defect for } \Delta \text{ if this exists} \\ \Delta & \text{otherwise} \end{cases}$$

Clearly, by  $\mathbf{Ax}_{\Box}$ , the set  $\Delta^+$  is consistent when  $\Delta$  is and hence, if  $\Delta$  is an approximation, so is  $\Delta^+$ . To define the extension  $\Sigma'$  of  $\Sigma$ , assume  $\varphi_0, \varphi_1, \dots$  to be an enumeration of the formulas in  $\mathcal{L}_{\omega}$ , and define

$$\begin{aligned} \Sigma_0 &= \Sigma \\ \Sigma_{2n+1} &= \begin{cases} \Sigma_{2n} \cup \{\varphi_n\} & \text{if this is consistent} \\ \Sigma_{2n} \cup \{\neg \varphi_n\} & \text{else} \end{cases} \\ \Sigma_{2n+2} &= (\Sigma_{2n+1})^+ \\ \Sigma' &= \bigcup_{n \in \omega} \Sigma_n \end{aligned}$$

By construction,  $\Sigma'$  is a maximal  $\mathbf{K}(A, \pi, \rho, I)$ -consistent, witnessed  $\pi'$ -theory extending  $\Sigma$ .  $\square$

**Definition 10 (Canonical model).** Given  $A, \pi, \rho$  and  $I$ , we define the canonical model  $M^c = (W^c, R^c, I, V^c)$  for the logic  $\mathbf{K}(A, \pi, \rho, I)$  as follows, where  $\pi'$  is as in Lemma 2:

- $W^c = \{\Gamma \mid \Gamma \text{ is a maximal } \mathcal{L}_\omega\text{-consistent witnessed } \pi'\text{-theory}\}$
- $R_b^c \Gamma \Delta$  iff for all  $\varphi \in \mathcal{L}_\omega$  it holds that if  $[a]\varphi \in \Gamma$ , then  $\varphi \in \Delta$
- $I$  is given as a parameter of the logic
- $V_p^c = \{\Gamma \mid p \in \pi' \cap \Gamma\}$

**Lemma 3** (Existence Lemma). *Let  $\Gamma$  be a maximal  $\mathbf{K}(A, \pi', \rho, I)$ -consistent witnessed  $\pi'$ -theory, with  $\langle b \rangle \varphi \in \Gamma$ . Then there is a maximal  $\mathbf{K}(A, \pi', \rho, I)$ -consistent witnessed  $\pi'$ -theory  $\Delta$  with  $\Gamma R_b^c \Delta$ .*

**Proof.** Let  $\varphi_0, \varphi_1, \dots$  be an enumeration of all formulas of the form  $\langle s \rangle \neg \Box(\vec{a})$ . Define the following sequence  $Y_0, Y_1, \dots$  of sets of formulas, where  $y_n = \bigwedge_{\psi \in Y_n} \psi$ :

$$\begin{aligned} Y_0 &= \{\varphi\} \\ Y_{2n+1} &= \begin{cases} Y_{2n} \cup \{\varphi_n\} & \text{if } \langle b \rangle (y_n \wedge \varphi_n) \in \Gamma \\ \Sigma_{2n} \cup \{\neg \varphi_n\} & \text{else, i.e., if } \langle b \rangle (y_n \wedge \neg \varphi_n) \in \Gamma \end{cases} \\ Y_{2n+2} &= (Y_{2n+1})^0 \\ Y &= \bigcup_{n \in \omega} Y_n \cup \{\sigma \mid [b]\sigma \in \Gamma\} \end{aligned}$$

Here, the operation  $^0$  is defined as follows. First of all, if  $\neg \varphi_n$  is added to  $Y_{2n+1}$ , then  $Y_{2n+1}^0 = Y_{2n+1}$ . Next, suppose  $\varphi_n = \langle s \rangle \neg \Box(\vec{a})$  was added to  $Y_{2n+1}$ . This means that  $\langle b \rangle (y_n \wedge \langle s \rangle \neg \Box(\vec{a})) \in \Gamma$ , i.e.,  $\langle b, y_n, s \rangle \neg \Box(\vec{a}) \in \Gamma$ . Since  $\Gamma$  is witnessed, we have  $\langle b, y_n, s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p}) \in \Gamma$  for some  $\vec{p}$ , and we put  $Y_{2n+1}^0 = Y_{2n+1} \cup \{\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p})\}$ .

Clearly,  $Y$  is fully witnessed. We claim  $Y$  is consistent as well, for suppose not, that is, there is a finite set of formulas from  $Y$  that together imply  $\perp$ . Let  $k$  be the biggest index for which we need a formula from  $Y_k$  to derive this contradiction. Then we have  $(y_k \wedge \alpha) \rightarrow \perp$ , for some  $\alpha$  in  $\{\sigma \mid [b]\sigma \in \Gamma\}$ , but this contradicts the facts that  $\langle b \rangle y_k \wedge [b]\alpha \in \Gamma$  and that  $\Gamma$  is consistent.

Finally, take a maximal consistent set  $\Delta$  around  $Y$ : clearly this set is witnessed (since by construction of  $Y$ , for every formula of the form  $\langle s \rangle \neg \Box(\vec{a})$ , a decision whether this formula or its negation should be included, is already made for  $Y$ ). Obviously, we have  $R_b^c \Gamma \Delta$ .  $\square$

At this point, we should mention that although the definition of canonical model is similar to the standard definition in modal logic (blackburnml), the domain of the model of Definition 10 is not the set of *all* maximal consistent sets, rather, it is the set of all such sets that are *witnessed*. The consequences of this fact will be discussed in Section 3.3. To prove our coincidence lemma below, we will now assume that  $\theta_{\Box}(\vec{a}, \vec{p})$  is *locally good*, in the following sense.

**Definition 11.** Let the logic  $\mathbf{K}(A, \pi, \rho, I)$  be given,  $\theta_{\Box}(\vec{a}, \vec{p})$  be a modal formula,  $\Box(\vec{a}) \in \rho$  a relational atom,  $\Theta(\vec{a})(x)$  a first order formula and  $I : \rho \rightarrow \mathcal{L}^1(A, \mathcal{X})$ . Assume that  $\Box(\vec{a})$  and  $\theta_{\Box}(\vec{a}, \vec{p})$  are connected through the axiom **Ax $_{\Box}$**  and **R $_{\Box}$** .

1. We say that  $\theta_{\Box}(\vec{a}, \vec{p})$  is *locally good* for  $\Theta(\vec{a})(x)$  iff
  - (a)  $\theta_{\Box}(\vec{a}, \vec{p})$  locally defines  $\Theta(\vec{a})(x)$ , and
  - (b) Let  $M^c$  be any canonical model obtained as in Definition 10. Then if  $\Gamma \in W^c$  contains all instances of  $\theta_{\Box}(\vec{a}, \vec{p})$ , then  $\Theta(\vec{a})(\Gamma)$  holds in the model  $M^c$ .
2. We finally say that the four-tuple  $\langle \theta_{\Box}(\vec{a}, \vec{p}), \Box(\vec{a}), I, \Theta(\vec{a})(x) \rangle$  is in local harmony if
  - (a)  $\theta_{\Box}(\vec{a}, \vec{p})$  is locally good for  $\Theta(\vec{a})(x)$ , and
  - (b)  $I$  semantically links  $\theta_{\Box}(\vec{a}, \vec{p})$  and  $\Box(\vec{a})$  through  $\Phi(x)$ .

**Lemma 4** (Coincidence Lemma). *Let  $M^c$  be defined as in Definition 10 and suppose for every  $\Box \in \rho$  there are purely modal formulas  $\theta_{\Box}(\vec{a}, \vec{p})$  and a first order formula  $\Theta(\vec{a})(x)$  such that the tuple  $\langle \theta_{\Box}(\vec{a}, \vec{p}), \Box(\vec{a}), I, \Theta(\vec{a})(x) \rangle$  is in local harmony. Then*

$$\text{For all } \varphi \in \mathcal{L}(A, \pi, \rho), \Gamma \in W^c: M^c, \Gamma \models \varphi \text{ iff } \varphi \in \Gamma$$

**Proof.** For propositional variables this holds by definition, and the Boolean connectives are immediate. For  $\varphi = \Box(\vec{a})$ , suppose  $\Box(\vec{a}) \in \Gamma$ . By **Ax $_{\Box}$**  and **US**, all instances of  $\theta_{\Box}(\vec{a}, \vec{p})$  are in  $\Gamma$  as well. By items 1b and 2b of Definition 11, we then have  $I(\Box(\vec{a}))(\Gamma)$ , which by truth definition, yields  $M^c, \Gamma \models \Box(\vec{a})$ . Conversely, suppose  $\Box(\vec{a}) \notin \Gamma$ , then  $\neg \Box(\vec{a}) \in \Gamma$ , and, by construction of  $\Gamma$ , for some  $\vec{p}$ , we have  $\neg \theta_{\Box}(\vec{a}, \vec{p}) \in \Gamma$ . Since  $\theta_{\Box}(\vec{a}, \vec{p})$  is locally good, it follows from items 1a and 2b of Definition 11, that it is not the case that  $I(\Box(\vec{a}))(\Gamma)$  holds, i.e.,  $M^c, \Gamma \models \neg \Box(\vec{a})$ .

The modal case is standard: let  $\varphi \in \mathcal{L}(A, \pi, \rho)$  be  $\varphi = [b]\psi$ . Let  $\Gamma \in W^c$  and suppose  $M^c, \Gamma \models [b]\psi$ . We now claim that  $[b]\psi \in \Gamma$ , for suppose not, that is, suppose  $\langle b \rangle \neg \psi \in \Gamma$ . With Lemma 3 we find a maximal consistent witnessed  $\Delta$  for which  $R_b^c \Gamma \Delta$  and  $\neg \psi \in \Delta$ . By induction and the fact that  $\psi, \neg \psi \in \mathcal{L}(A, \pi, \rho)$  we also have  $M^c, \Delta \models \neg \psi$  which contradicts  $M^c, \Gamma \models [b]\psi$ . So  $\{\alpha \mid [b]\alpha \in \Gamma\} \cup \{\neg \psi\}$  is not  $\mathcal{L}$ -consistent and we have  $\{\alpha \mid [b]\alpha \in \Gamma\} \vdash \psi$  and hence  $\Gamma \vdash [b]\psi$  and

consequently  $[b]\psi \in \Gamma$ . For the other direction, suppose  $[b]\psi \in \Gamma$ . Take an arbitrary  $\Delta \in W^c$  with  $\psi \in \Delta$ . By induction,  $M^c, \Delta \models \psi$ . Since  $\Delta$  was arbitrary, we have  $R_b^c \Gamma \Delta$ .  $\square$

**Theorem 2.** Let the logic  $\mathbf{K}(A, \pi, \rho, I)$  be given as in Definition 7, and suppose for every  $\Box \in \rho$  there are purely modal formulas  $\theta_{\Box}(\vec{a}, \vec{p})$  and a first order formula  $\Theta(\vec{a})(x)$  such that the tuple  $\langle \theta_{\Box}(\vec{a}, \vec{p}), \Box(\vec{a}), I, \Theta(\vec{a})(x) \rangle$  is in local harmony. Then  $\mathbf{K}(A, \pi, \rho, I)$  is sound and complete with respect to the class of  $\mathcal{K}(A, \pi, \rho, I)$  models.

**Proof.** Soundness is dealt with in Theorem 1. For completeness, let  $\varphi \in \mathcal{L}(A, \pi, \rho)$  be such that  $\mathbf{K}(A, \pi, \rho, I) \not\models \neg\varphi$ . By Lemma 2,  $\{\varphi\}$  can be extended to a maximal  $\mathcal{L}_\omega$ -consistent witnessed theory  $\Gamma$ . By Lemma 4, we then have  $M^c, \Gamma \models \varphi$ , and hence  $\not\models \neg\varphi$  which completes the proof.  $\square$

### 3.3. Discussion and an extension of the completeness proof

In a nutshell, our completeness proof builds a model for maximal witnessed theories  $\Gamma$ . Every consistent formula  $\varphi$  can be embedded in such a theory, and the construction makes sure that for every  $\langle s \rangle \neg \Box(\vec{a}) \in \Gamma$ , there are witnessing atoms  $\vec{p}$  such that  $\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p}) \in \Gamma$ . Our canonical model hence does not contain all maximal consistent sets, but only those that are fully witnessed. Such a model is called a *thinned out canonical model* in [12].

In modal logics that extend  $\mathbf{K}$  with an axiom  $\varphi$ , the completeness proof for  $\mathbf{K} + \varphi$  typically builds the canonical model from all maximal consistent sets, and this construction works for axioms  $\varphi$  that are *canonical* for some first order property  $\Phi(x)$  (see also [3, Chapter 5] for a further discussion). Completeness of  $\mathbf{K} + \varphi$  is then obtained with respect to those frames of  $\mathbf{K}$  that moreover satisfy  $\Phi$ .

For our set-up, rather than canonicity, we require a property given in Definition 11. Note that our frames do not globally inherit the first order property  $\Theta(\vec{a})$ , but instead, at every point  $\Gamma$  in the canonical model where  $\theta_{\Box}(\vec{a}, \vec{p})$  holds, we also have the property  $\Theta(\vec{a})$ .

It appears<sup>1</sup> there is a rich modal literature on proving completeness when having a thinned out canonical model like ours. A general way to prove completeness for such logics is presented in [12] (which in turn, is based on reports that appear as the first chapter in [11]). We will not present the details here: it would involve the notion of *general frames*. Our proof shows that for the kind of model we define, the notion of  $\theta_{\Box}(\vec{a}, \vec{p})$  being *locally good* (Definition 11) works to prove completeness. It might be cumbersome to check this notion for every  $\theta_{\Box}(\vec{a}, \vec{p})$  separately, but luckily enough, from [12,11], it follows that our construction works for a class of formulas  $\theta_{\Box}(\vec{a}, \vec{p})$  that are called *locally r-persistent* (again, this notion is defined in terms of general frames, we refer to [12,11,30]). For an example of a formula that is canonical but not r-persistent, see [30].

A systematic characterisation of (locally) r-persistent formulas is still lacking, but [22, Theorem 2.4.7] ensures that shallow formulas, i.e., modal formulas in which every occurrence of a propositional variable is in the scope of at most one modal operator are r-persistent. Moreover, [12, p. 182] mentions that ‘many well-known formula axiomatising natural frame conditions like reflexivity, symmetry, transitivity, linearity etc. are locally r-consistent’.

So the fact that r-persistence of  $\theta_{\Box}(\vec{a}, \vec{p})$  implies that  $\theta_{\Box}(\vec{a}, \vec{p})$  is locally good for some first order  $\Theta(\vec{a})(x)$  implies that our completeness proof works whenever  $\theta_{\Box}(\vec{a}, \vec{p})$  is r-persistent. It moreover guarantees the following ‘modular’ completeness result.

**Theorem 3.** Consider the logic  $\mathbf{K}(A, \pi, \rho, I)$  and add a number of r-persistent formulas  $\varphi_1, \dots, \varphi_n$  as axioms to it: call the result  $\mathbf{K}'$ . Suppose  $\varphi_i$  locally defines  $\Phi_i$ . Then  $\mathbf{K}'$  is complete with respect to frames that satisfy  $\Phi_i$  ( $i \leq n$ ).

**Proof.** For every axiom  $\varphi_i$ , introduce a new relational atom  $\Box_{\varphi_i}$  which is added as an axiom to the logic, and also add an axiom and inference rule that axiomatically link  $\varphi_i$  and  $\Box_{\varphi_i}$ . Our completeness proof works when having several relational atoms  $\Box$  associated with a formula  $\theta_{\Box}(\vec{a}, \vec{p})$  that is locally good for  $I(\Box)$ . Since r-persistence of  $\varphi_i$  implies being locally good for  $\Phi_i$ , we obtain a model such that every set  $\Gamma$  that contains  $\Box_{\varphi_i}$  satisfies  $\Phi_i(\Gamma)$ , and moreover, since  $\Box_{\varphi_i}$  is added as an axiom, the whole model satisfies  $\Phi$ . This completes the proof.  $\square$

We will see an immediate application of Theorem 3 in the next section.

Our inference rule  $\mathbf{R}_{\Box}$  is reminiscent of an inference rule for irreflexivity [8]. This rule for irreflexivity triggered a flurry of papers on studying similar ‘unorthodox rules’. We only mention here [11,30,12], which establish some general completeness results, technically similar to the one presented here. The paper [30] calls such rules ‘non- $\xi$  rules’ and [12] calls them ‘context-dependent rules’, because of the similarity between such rules and context-dependent rules in generating grammars. Also, our use of *pseudo-modalities* has ancestors: (often in combination with the kind of ‘unorthodox rules’ above), they are used in [12,30] as ‘universal forms’, in [9] as ‘admissible forms’ and were already present as a prototype in [8]. They also play a similar role as does the ‘pasting’ of a subformula next to an occurrence of a specified one in [6] to prove the completeness of the  $D$  (difference) operator.

<sup>1</sup> We are indebted to one of the reviewers to point this out to us.

However, it is important to stress that the approaches mentioned aim to axiomatise *global* properties. As far as we know, the work presented in this paper is a first general approach to *local* properties in models. Although [12] also pays attention to *local* issues, all the frames obtained there are defined globally. Theorem 4.1 in [12] presents both a generalisation (it is not just about rules with relational atoms) and a specialisation (it is not about *I*-frames) of our Theorem 2. In our terminology, [12, Theorem 4.1] supposes that we have a formula  $\varphi$  that locally defines  $\alpha(x)$ , and a formula  $\psi$  which is locally *r*-persistent and locally defining  $\beta(x)$ . It then assumes an inference rule like  $\mathbf{R}_{\Box}$ , where  $\theta_{\Box}(\vec{a}, \vec{p})$  is replaced by  $\varphi$  and  $\Box(\vec{a})$  by  $\psi$ . If such a rule is added to a logic  $\mathbf{L}$  which only has *r*-persistent axioms, then the resulting logic is sound and complete with respect to frames of  $\mathbf{L}$  that moreover satisfy  $\forall x(\alpha(x) \rightarrow \beta(x))$ . In our case, the  $\alpha$  and  $\beta$  are equivalent, so that indeed we keep the frames of the underlying logic  $\mathbf{L}$ . In other words, our logic does not demand anything about a frame globally, it only requires that if locally all instances of  $\theta_{\Box}(\vec{a}, \vec{p})$  are true, the ‘corresponding’  $I(\Box(\vec{a}))$  must hold as well.

### 3.4. Outlook

We have so far assumed that the only properties of  $\Box(\vec{a})$  are those specified by the axiom  $\mathbf{A}_{\Box}$  and rule  $\mathbf{R}_{\Box}$ . However, one can add other connections between  $\Box(\vec{a})$  and modal formulas, or between different  $\Box_1(\vec{a}_1)$  and  $\Box_2(\vec{a}_2)$  atoms. For instance, in an epistemic logic, one could add the axiom scheme

$$\text{Refl}(a, a) \rightarrow \text{Trans}(a, a, a) \quad (5)$$

Adding an axiom like (5) for an agent  $a$  has the effect that whenever  $a$ ’s knowledge is veridical,  $a$  is also positively introspective. I.e., we would have, semantically, that whenever  $M, s \models K_a \varphi \rightarrow \varphi$ , for all  $\varphi$ , then also  $M, s \models K_a \varphi \rightarrow K_a K_a \varphi$ , for all  $\varphi$ . This again is a property that cannot be expressed in standard, ‘global’ modal logic. As a second example, in an epistemic temporal modal logic, one could add an axiom like

$$\text{Trans}(a, a, a) \rightarrow \Diamond(\text{Trans}(a, a, a) \wedge \text{Eucl}(a, a, a)) \quad (6)$$

saying that whenever agent  $a$  is positively introspective, he will eventually also become negatively introspective. As a third example, a simple axiom like

$$\text{Ser}(a) \rightarrow \text{Ser}(b) \quad (7)$$

in a doxastic setting would mean that whenever  $a$ ’s beliefs are consistent, those of  $b$  must be consistent as well.

It is possible to view some standard results in modal logic concerning completeness of modal systems as special cases of our local logic. If the conditions of Theorem 2 are satisfied, and one adds a  $\Box(\vec{a})$  as an *axiom*, one immediately gets completeness with respect to the class of models that satisfy  $I(\theta_{\Box}(\vec{a}, \vec{p}))$ . For instance, in a logic with axioms and rules for  $\text{Refl}(a)$ , adding  $\text{Refl}(a)$  itself as an axiom gives a modal system that is sound and complete with respect to the class of reflexive Kripke models! Of course, this amounts to the same thing as adding  $\theta_{\Box}(\vec{a}, \vec{p})$ , as is directly clear from rule  $\mathbf{R}$  (take  $\alpha = \perp$ ).

Finally, it is important to realise that, although we presented the axioms for the underlying logic (the formulas  $\varphi_i$  that we assumed to be canonical) and the relational atoms as two independent layers, the interaction properties between the modalities and the relational atoms may be automatically ‘imported’. For the case of epistemic logic  $\mathbf{S5}$  with at least two agents and the  $\text{Sup}(a, b)$  atom for instance, one can derive that certain  $\text{Sup}(a, b)$  statements cannot go unnoticed by the agents! This will now be discussed in the next section.

## 4. Comparative (epistemic) logic

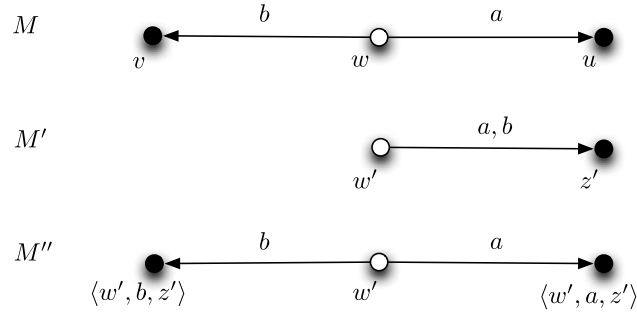
We will now focus on an example of a modal logic, called *comparative logic* ( $\mathbf{CL}$ ), which has one type of relational atom  $\Box$ , namely  $\text{Sup}$ . From now on, we will write  $a \succcurlyeq b$  for  $\text{Sup}(a, b)$ , and  $b \preccurlyeq a$  will denote the same. To motivate our notation, note that

$$M, w \models a \succcurlyeq b \quad \text{iff} \quad R_a(w) \supseteq R_b(w) \quad \text{iff} \quad M, w \models b \preccurlyeq a \quad \text{iff} \quad M, w \models R_b(w) \subseteq R_a(w)$$

We also write  $a \not\succcurlyeq b$  for  $\neg(a \succcurlyeq b)$ . Moreover, we will write both  $a \succ b$  and  $b \prec a$  for  $(a \succcurlyeq b) \wedge \neg(b \succcurlyeq a)$ . So  $a \preccurlyeq b$  means ‘every  $a$ -successor is a  $b$ -successor’, and  $a \prec b$  means ‘every  $a$ -successor is a  $b$ -successor, and some  $b$ -successor is not an  $a$ -successor’. In the second part of this section, we will then interpret our formulas on multi-agent  $\mathbf{S5}$ -models (‘epistemic’ models). So then  $a \succcurlyeq b$  stands for ‘agent  $a$  considers a larger set of worlds possible than agent  $b$ ’, informally ‘ $b$  knows at least as much as  $a$ ’. We will start by adding the atom to the logic  $\mathbf{K}(A, \pi, \rho, I)$ . This is the logic  $\mathbf{CL}$ : *comparative logic*. This is not yet an *epistemic* logic.

### 4.1. Comparative logic

**Definition 12** (*Comparative logic*). *Comparative logic*,  $\mathbf{CL}(A, \pi)$ , is the logic that is obtained from  $\mathbf{K}(A, \pi, \rho, I)$  by choosing as  $\rho$  the singleton consisting of relational atom  $\Box = \succcurlyeq$  (i.e., we have all  $a \succcurlyeq b$  for  $a, b \in A$ , in the language); where we stipulate



**Fig. 3.** The countermodels in Example 3 and Lemma 5. The atom  $p$  is true only in the black worlds.

that  $I(a \succ b) = \forall xy(R_bxy \Rightarrow R_axy)$  and, finally, that  $\theta_{\Box}(a, b, p) = [a]p \rightarrow [b]p$ . In  $\mathbf{CL}(A, \pi)$ , the atom  $a \succ b$  and  $[a]p \rightarrow [b]p$  are axiomatically linked through the axiom

$$\mathbf{Ax}_{\succ} = a \succ b \rightarrow ([a]p \rightarrow [b]p)$$

and inference rule  $\mathbf{R}_{\succ}$ :

$$\mathbf{R}_{\succ} = \text{From } [s]([a]p \rightarrow [b]p) \text{ infer } [s](a \succ b) (p \text{ not in } s)$$

which is equivalent to:

$$\mathbf{R}_{\succ} = \text{From } \langle s \rangle \neg([a]p \rightarrow [b]p) \rightarrow \varphi \text{ infer } \langle s \rangle \neg(a \succ b) \rightarrow \varphi (p \text{ not in } \varphi \text{ or } s)$$

We write  $\vdash_{\mathbf{CL}}$  for derivability in  $\mathbf{CL}(A, \pi)$ . The language  $\mathcal{CL}(A, \pi)$  of comparative logic is defined as  $\mathcal{L}(A, \pi, \{\succ\})$ .

**Corollary 1.**  $\mathbf{CL}(A, \pi)$  is sound and complete with respect to  $\mathcal{K}(A, \pi, \rho, I)$

**Proof.** This follows directly from Theorem 2 on page 143.  $\square$

In this weakest modal logic with the symbol  $\succ$ , we can already derive ‘expected’ properties of  $\succ$ :

**Proposition 1.** The following hold:

1.  $\vdash_{\mathbf{CL}} a \succ a$
2.  $\vdash_{\mathbf{CL}} (a \succ b \wedge b \succ c) \rightarrow a \succ c$
3.  $\vdash_{\mathbf{CL}} a \succ b \rightarrow \langle a \rangle \top$

**Proof.** 1. Since  $[a]p \rightarrow [a]p$  is a theorem, we have  $\vdash_{\mathbf{CL}} \neg([a]p \rightarrow [a]p) \rightarrow \perp$  (where  $\perp = (q \wedge \neg q)$ ). Applying the rule  $\mathbf{R}_{\succ}$  to this gives  $\vdash_{\mathbf{CL}} \neg(a \succ a) \rightarrow \perp$ . The latter is equivalent to  $\vdash_{\mathbf{CL}} a \succ a$ .

2. Using  $\mathbf{Ax}_{\succ}$ , we have  $\vdash_{\mathbf{CL}} (a \succ b \wedge b \succ c) \rightarrow (([a]p \rightarrow [b]p) \wedge ([b]p \rightarrow [c]p))$ . This implies  $\vdash_{\mathbf{CL}} (a \succ b \wedge b \succ c) \rightarrow ([a]p \rightarrow [c]p)$ . We can now use rule  $\mathbf{R}^2$  (with  $s$  the empty sequence) to conclude  $\vdash_{\mathbf{CL}} (a \succ b \wedge b \succ c) \rightarrow a \succ c$ .

3. By definition,  $a \succ b$  equals  $(a \succ b) \wedge \neg(b \succ a)$ . We also have  $\vdash_{\mathbf{CL}} \neg([b]p \rightarrow [a]p) \rightarrow \langle a \rangle \top$ . Using  $\mathbf{R}_{\succ}$  we then get  $\vdash_{\mathbf{CL}} \neg(b \succ a) \rightarrow \langle a \rangle \top$ . Using the definition above, we then get  $\vdash_{\mathbf{CL}} a \succ b \rightarrow \langle a \rangle \top$ .  $\square$

Our next example reminds the reader that  $a \succ b$  and  $[a]\varphi \rightarrow [b]\varphi$  are not the same.

**Example 3.** Suppose  $\Gamma$  is the set of formulas

$$\Gamma = \{[a]\varphi \rightarrow [b]\varphi \mid \varphi \in \mathcal{L}\} \cup \{a \not\succ b\}$$

Then it is well possible that  $\Gamma$  is satisfiable, and, by soundness, consistent. For satisfiability, take the following model  $M = \langle W, R, I, V \rangle$ , with  $W = \{w, u, v\}$ ,  $R_a = \{(w, u)\}$  and  $R_b = \{(w, v)\}$ ,  $I(a \succ b)(x) = \forall y(R_bxy \Rightarrow R_axy)$  and, finally,  $V_p = \{u, v\}$  for all  $p$ . (See Fig. 3.) In this model, we have  $M, u \models \varphi$  iff  $M, v \models \varphi$  for all  $\varphi$ , and hence we have  $M, w \models [a]\varphi \rightarrow [b]\varphi$ , and yet we do not have  $R_b(w) \subseteq R_a(w)$ , i.e.  $M, w \models a \not\succ b$ .

To further emphasise the non-standard behaviour of our modal language, we state two more negative (and one positive) results. First let us briefly revisit some modal semantic notions.

**Definition 13.** Given two models  $M = \langle W, R, V \rangle$  and  $M' = \langle W', R', V' \rangle$ , a relation  $\mathcal{R} \subseteq W \times W'$  is called a *bisimulation* if the following holds: ('atomic') for all  $p \in P$ , if  $\mathcal{R}ww'$  then  $w \in V(p)$  iff  $w' \in V'(p)$  ('forth') if  $\mathcal{R}ww'$  and if for some  $v \in W$  and some  $a \in A$  one has  $R_a wv$ , then there is a  $v' \in W'$  such that  $R'_a w'v'$  and  $\mathcal{R}vv'$  and, finally ('back') if  $\mathcal{R}ww'$  and if for some  $v' \in W'$  and some  $a \in A$  one has  $R'_a w'v'$ , then there is a  $v \in W$  such that  $R_a wv$  and  $\mathcal{R}vv'$ . If there is a bisimulation between  $M$  and  $M'$  with  $\mathcal{R}ww'$ , we write  $M, w \cong M', w'$ .

A special case of a bisimulation is obtained by unraveling a model  $(M, w)$  into a model  $(M', w)$  as follows. Given  $W$  and a set of agents  $A$  let  $W'$  be all the finite paths in  $M$  from  $w$ , i.e., states  $w'$  in  $W'$  are of the form  $w' = \langle w_1, a_1, w_2, a_2, \dots, w_n, a_n w_{n+1} \rangle$  such  $w_1 = w$  and for all  $i \leq n$ , in  $M$  one has  $R_{a_i} w_i w_{i+1}$ . Let  $lst(w') = lst(\langle w_1, a_1, w_2, a_2, \dots, w_n, a_n w_{n+1} \rangle) = w_{n+1}$  (i.e.,  $lst$  selects the last member of a finite list). Put  $w' \in V'(p)$  iff  $lst(w') \in V(p)$ , and  $R'_a w'v'$  if  $v' = \langle w_1, a_1, w_2, a_2, \dots, w_n, a_n w_{n+1}, a, u \rangle$  for some  $u \in W$ .

If there are no relational atoms we have a standard multimodal logical language: for  $\mathcal{L}(A, \pi, \emptyset)$  we write  $\mathcal{L}(A, \pi)$ . In other words, compared to  $CL(A, \pi) = \mathcal{L}(A, \pi, \{\succ\})$ ,  $\mathcal{L}(A, \pi)$  is the language without the  $a \succ b$  formulas.

**Lemma 5.**

1. *Bisimulations preserve  $\mathcal{L}(A, \pi)$*  [3, p. 66]:  
If  $M, w \cong M', w'$  then for all  $\varphi \in \mathcal{L}(A, \pi)$ :  $M, w \models \varphi$  iff  $M', w' \models \varphi$ .
2. *Unravelings preserve  $\mathcal{L}(A, \pi)$*  [3, p. 63]:  
If  $M', w$  is an unraveling of  $M, w$  then for all  $\varphi \in \mathcal{L}(A, \pi)$ :  $M, w \models \varphi$  iff  $M', w \models \varphi$ .
3. *Bisimulations do not preserve  $CL(A, \pi)$*
4. *Unravellings do not preserve  $CL(A, \pi)$*

**Proof.** We only show items 3 and 4. Take the model  $M' = \langle W', R', V' \rangle$  such that  $W' = \{w', z'\}$  and  $R'_a = R'_b = \{(w', z')\}$ . (See Fig. 3.) Then for  $M$  defined in Example 3, it is clear that  $M, w \cong M', w'$ , yet  $M', w' \models b \succ a$  while  $M, w \models \neg(b \succ a)$ . For item 4, consider the unraveling  $M'' = \langle W'', R'', V'' \rangle$  of  $M'$ , where  $W'' = \{w', \langle w', a, z' \rangle, \langle w', b, z' \rangle\}$ . (See Fig. 3.) It is easily verified that  $M', w' \models a \succ b \wedge b \succ a$  while  $M'', w' \models \neg(a \succ b) \wedge \neg(b \succ a)$ . The latter in fact shows a validity for any unravelled model (for  $a \neq b$ ).  $\square$

To show that  $CL(A, \pi)$  is not completely misbehaved, we show that there are modified kinds of bisimulation and unraveling that *do* preserve the language. The idea is simple: instead of looking at individual steps we look at complete coalitions for which two states are accessible.

**Definition 14.** Let for a coalition  $C \subseteq A$ , relation  $R_C$  be such that  $R_C wv$  iff for all  $i \in A$ : ( $R_i wv$  iff  $i \in C$ ). Note the second occurrence of 'iff' in this definition: it follows that for every  $w$  and  $v$  there is exactly one coalition  $C \subseteq A$  for which  $R_C wv$ .

Given two models  $M = \langle W, R, V \rangle$  and  $M' = \langle W', R', V' \rangle$ , a relation  $\mathcal{R} \subseteq W \times W'$  is called a *coalitional bisimulation* if the following holds: ('atomic') for all  $p \in \pi$ , if  $\mathcal{R}ww'$  then  $w \in V(p)$  iff  $w' \in V'(p)$  ('forth') if  $\mathcal{R}ww'$  and if for some  $v \in W$  and some  $C \subseteq A$  one has  $R_C wv$ , there is a  $v' \in W'$  such that  $R'_C w'v'$  and  $\mathcal{R}vv'$  and, finally ('back') if  $\mathcal{R}ww'$  and if for some  $v' \in W'$  and some  $C \in A$  one has  $R'_C w'v'$ , there is a  $v \in W$  such that  $R_C wv$  and  $\mathcal{R}vv'$ . If there is a coalitional bisimulation between  $M$  and  $M'$  with  $\mathcal{R}ww'$ , we write  $M, w \cong_{coal} M', w'$ .

Let  $M, w$  with  $M = \langle W, R, V \rangle$  be given. A *coalitional unraveling* of  $M$  is a model  $M', w$  with  $M' = \langle W', R', V' \rangle$ , and where  $W'$  consists of all paths  $\langle w_1, C_1, w_2, \dots, w_n, C_n, w_{n+1} \rangle$  such that  $w_1 = w$  and for all  $i \leq n$ , one has  $R_{C_i} w_i w_{i+1}$ .  $V'$  is defined as in the case for (ordinary) unravelings, and  $R'_C w'v'$  if  $v' = \langle w_1, C_1, w_2, C_2, \dots, w_n, C_n w_{n+1}, C, u \rangle$  for some  $u \in W$  and  $C \subseteq A$ .

Coalitional unravelings  $(M', w)$  of  $(M, w)$  respect access for *coalitions* of agents, in the following sense: for all  $C$ ,  $R'_C w'v'$  iff  $R_C lst(w')lst(v')$ . The proof of the following is now a straightforward extension of the proofs of items 1 and 2 of Lemma 5 and therefore omitted:

**Theorem 4 (Preservation).** *We have the following.*

1. If  $M, w \cong_{coal} M', w'$  then for all  $\varphi \in CL(A, \pi)$ :  $M, w \models \varphi$  iff  $M', w' \models \varphi$ .
2. If  $M', w$  is a coalitional unraveling of  $M, w$  then for all  $\varphi \in CL(A, \pi)$ :  $M, w \models \varphi$  iff  $M', w \models \varphi$ .

**Proposition 2.** *A coalitional bisimulation is a bisimulation: if  $M, w \cong_{coal} M', w'$  then  $M, w \cong M', w'$ .*

**Proof.** A proof is needed, because the back and forth requirements for the bisimulation relation are not special cases of those for coalitional bisimulation, as  $R_a$  is not the same as (in coalitional bisimulations)  $R_{\{a\}}$ : the latter links states with only an  $a$ -arrow between them, the former links states with *at least* an  $a$ -arrow between them.



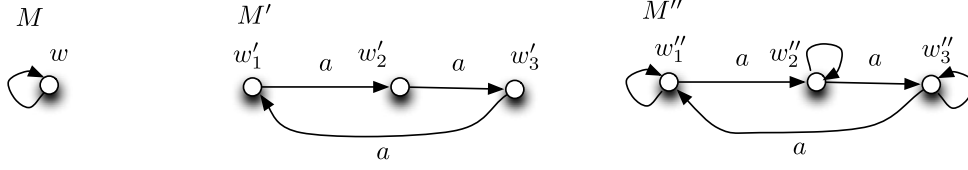


Fig. 4. Three bisimilar models  $M, w, M', w'_1$  and  $M'', w''_1$ .

Assume  $\mathcal{R}_{coal}: M, w \cong_{coal} M', w'$ . We show that also  $\mathcal{R}_{coal}: M, w \cong M', w'$ . ‘Atoms’ is trivial. ‘Forth’: let  $a \in A, u \in M$ , and  $u' \in M'$ . Assume there is a  $v \in M$  such that  $R_a uv$  and  $\mathcal{R}_{coal} uv'$ . Agent  $a$  must occur in some coalition  $C$  for which  $R_C uv$  (there is a  $C$  such that  $\{a\} \subseteq C \subseteq A$  and  $R_C$  non-empty). From  $R_C uv$  and  $\mathcal{R}_{coal} uv'$  follows that there is a  $v' \in M'$  such that  $R'_C u'v'$  and  $\mathcal{R}_{coal} vv'$ . From  $R'_C u'v'$  and  $a \in C$  follows  $R'_a u'v'$ . ‘Back’: similar.  $\square$

The notion of coalitional bisimulation is in fact too strong to characterise preservation, i.e., the converse direction of item 1 of Theorem 4 does not hold. To see this, take the two models  $M, w_1$  and  $M', w'_1$  of Fig. 11 (in fact, all accessibility relations are equivalence relations: the formal definition of the models is in the proof of Theorem 8). We will argue in that proof that  $M, w_1$  and  $M', w'_1$  satisfy the same formulas, but it is easy to see that they are not coalition bisimilar: from  $w_1$ , there is an  $\{a\}$ -successor that satisfies  $\neg p$ , but this is not true for  $w'_1$ , nor for  $w'_2$ . But this also suggest a trivial relaxation of our requirements for a coalitional bisimulation: in order for a state  $M', w'$  to have the same theory as  $M, w_1$  in Fig. 11, rather than requiring that  $M', w'$  needs to have an  $\{a\}$ -successor to a  $\neg p$  world, we should require that (i)  $M', w'$  is bisimilar with  $M, w$  in the sense of Definition 13, and (ii), the formula  $b \preceq a$  is true in  $M', w'$  (and it should agree with  $M, w_1$  on all relevant  $\preceq$ -formulas).

**Definition 15.** Let the models  $M = \langle W, R, I, V \rangle$  and  $M' = \langle W', R', I', V' \rangle$  be two models for the same agents  $A$ , propositional variables  $\pi$  and relational atoms  $\rho$ . A relation  $\mathcal{R} \subseteq W \times W'$  is called an *extended bisimulation* if

- $\mathcal{R}$  is a bisimulation (in the sense of Definition 13).
- If for  $w \in W, w' \in W'$  we have  $\mathcal{R}ww'$ , then  $I(\Box(\vec{a}))(w)$  iff  $I'(\Box(\vec{a}))(w')$  for all  $\Box(\vec{a}) \in \rho$ .

An extended bisimulation is moreover called *natural* if it also satisfies

- If for  $w \in W, w' \in W'$  we have  $\mathcal{R}ww'$ , then  $I(\Box(\vec{a}))(w) = I'(\Box(\vec{a}))(w')$  for all  $\Box(\vec{a}) \in \rho$ .

As an example, consider the models  $M = \langle W, R, I, V \rangle$ ,  $M' = \langle W', R', I', V' \rangle$  and  $M'' = \langle W'', R'', I'', V'' \rangle$  of Fig. 4. Let  $V, V'$  and  $V''$  be such that exactly the same propositional variables are true in all worlds. Suppose there is only one relational atom  $\Box(a)$  in the language, and  $I(\Box(a)(x)) = R_a xx$ ,  $I'(\Box(a)(x)) = \exists! y R_a xy$  and  $I''(\Box(a)(x)) = R_a xx$ . Then  $\mathcal{R} = W \times W'$  is an extend bisimulation that is not natural (in  $M$ ,  $\Box(a)$  is interpreted as  $R_a$  being reflexive, in  $M'$  as  $R'_a$  being functional), while  $\mathcal{R}' = W \times W''$  represents a natural extended bisimulation (in both models,  $\Box(a)$  is interpreted as reflexivity). Note that for  $I'(\Box(a)(x))$  we could even have taken  $\neg I(\Box(a)(x))$ . In particular, if  $N'$  is an unraveling a model  $N$  with a reflexive or transitive relation  $R_a$ , then the two models bisimulate each other non-naturally by taking for instance  $I(\Box(a)$  to be  $Ref(a)$  or  $Trans(a)$ , and  $I'$  its negation. To sum up, in an extended bisimulation, bisimulating worlds agree also on the *truth* of relational atoms, while in a natural extended bisimulation, bisimulating worlds moreover agree on the *meaning* of such atoms.

The following is an easy extension of well-known results in modal logic. Note that it is about general languages  $\mathcal{L}(A, \pi, \rho)$ , not just about the comparative language of  $\mathbf{CL}(A, \pi)$ .

**Lemma 6.**

1. *Extended bisimulations preserve  $\mathcal{L}(A, \pi, \rho)$ , i.e., if there is an extended bisimulation between  $M, w$  and  $M', w'$ , then for all  $\varphi \in \mathcal{L}(A, \pi, \rho)$ ,  $M, w \models \varphi$  iff  $M', w' \models \varphi$ .*
2. *For finite models, we also have the converse: if two finite models agree on formulas form  $\mathcal{L}(A, \pi, \rho)$ , then there exists an extended bisimulation between them.*

**Proof.** Similar to the modal case (cf. [3]), when one treats relational atoms as propositional variables.  $\square$

#### 4.2. Comparative epistemic logic

We will now specialise our case study with  $\Box(a, b) = a \succ b$  to epistemic logic. We will write  $K_a \varphi$  (rather than  $[a]\varphi$ ) for ‘agent  $a$  knows  $\varphi$ ’. Moreover,  $M_a \varphi$  will be short for  $\neg K_a \neg \varphi$ .

So our language  $\mathcal{L}(\{K_a\}_{a \in A}, \pi, \{a \succ b\}_{a, b \in A})$ , or  $\mathcal{L}_{\mathbf{CEL}}$  for short, is defined as follows, where  $\pi$  is as described earlier.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_a\varphi \mid a \succ b$$

**Definition 16** (Comparative epistemic logic). Let  $\mathbf{CEL}(A, \pi)$ , *comparative epistemic logic*, be the logic that is obtained from  $\mathbf{CL}(A, \pi)$  by adding the following three knowledge axioms.

- T**  $K_a\varphi \rightarrow \varphi$
- 4**  $K_a\varphi \rightarrow K_a K_a\varphi$
- 5**  $\neg K_a\varphi \rightarrow K_a \neg K_a\varphi$

Moreover, since we assume  $\rho = \{a \succ b \mid a, b \in A\}$  we have the following specialised axiom and rule (see Table 4 on page 139; we recall that  $s$  is a pseudo-modality):

- Ax <sub>$\succ$</sub>**   $a \succ b \rightarrow (K_a p \rightarrow K_b p)$
- R <sub>$\succ$</sub>**  From  $\langle s \rangle (K_a p \wedge \neg K_b p) \rightarrow \varphi$  infer  $\langle s \rangle \neg (a \succ b \rightarrow \varphi)$

As in **CL**, we assume in **CEL** that  $I(a \succ b) = \forall y (R_b x y \Rightarrow R_a x y)$  and, finally, that  $\theta_{\square}(a, b, p) = K_a p \rightarrow K_b p$ . So, in  $\mathbf{CEL}(A, \pi)$ , we have axioms that guarantee that  $a \succ b$  and  $K_a p \rightarrow K_b p$  are axiomatically linked. We will write  $\vdash_{\mathbf{CEL}}$  for derivability in  $\mathbf{CEL}(A, \pi)$ .

Models for  $\mathbf{CEL}(A, \pi)$  will be ordinary S5-models, i.e., models of the kind  $M = \langle W, R, I, V \rangle$  with the constraint that for each  $a \in A$ ,  $R_a$  is an equivalence relation. Call this set of models  $S5(A, \pi, \rho, I)$ .

**Theorem 5.**  $\mathbf{CEL}(A, \pi)$  is sound and complete with respect to  $S5(A, \pi, \rho, I)$ .

**Proof.** This follows directly from Theorem 3 and the observation that axiom **T** is r-persistent and defines reflexivity, axiom **4** if r-persistent and defines transitivity, and **5** r-persistent and defines Euclidicity (for the r-persistence argument, see [12]). Moreover, a relation has those three properties iff it is an equivalence relation. Finally, **Ax <sub>$\succ$</sub>**  is also r-persistent, since it is shallow [22].  $\square$

The technical results in the previous sections suggest that the infinite scheme

$$\bigwedge_{\varphi \in \mathcal{L}} (K_a \varphi \rightarrow K_b \varphi) \tag{8}$$

is captured by the formula  $a \succ b$ . However, we have also seen that, would we allow for infinite conjunctions, then although we would have  $\vdash (a \succ b) \rightarrow \bigwedge_{\varphi \in \mathcal{L}} (K_a \varphi \rightarrow K_b \varphi)$ , this implication can in general not be reversed.

Related to this, it is worth noting that the notion of ‘knowing more than’ cannot be captured in the language of **CEL**. Although one might suspect that ‘ $b$  knows more than  $a$ ’ is captured by  $a \succ b$ , the latter only says that *every state considered possible by  $b$  is also considered possible by  $a$  and, moreover, some states are considered possible by  $a$  and not by  $b$* . However, these latter states can all be bisimilar to states that both agents consider possible, in which case both agents would know the same.

The following theorem summarises some of our discussion so far.

**Theorem 6.** Let  $(M, w)$  be a pointed **CEL** model. Let  $\Psi(w)$  denote that in  $w$ , agent  $b$  knows at least what  $a$  knows:

$$\Psi(w) : \forall \varphi ((M, w) \models K_a \varphi \rightarrow K_b \varphi)$$

Furthermore, let  $\Phi(w)$  denote that at  $w$ , agent  $b$  knows strictly more than  $a$ :

$$\Phi(w) : \forall \varphi ((M, w) \models K_a \varphi \rightarrow K_b \varphi) \ \& \ \exists \psi ((M, w) \models K_b \psi \wedge \neg K_a \psi)$$

1. If  $(M, w) \models a \succ b$ , then  $\Psi(w)$ ;
2. The converse of item 1 does not hold in general;
3.  $(M, w) \models a \succ b$  does not imply  $\Phi(w)$  in general;
4.  $\Phi(w)$  does not in general imply  $(M, w) \models a \succ b$ .

**Proof.** All items, except item 1, require a counter-example.

1. This follows from **Ax <sub>$\succ$</sub>**  and soundness of **CEL**.

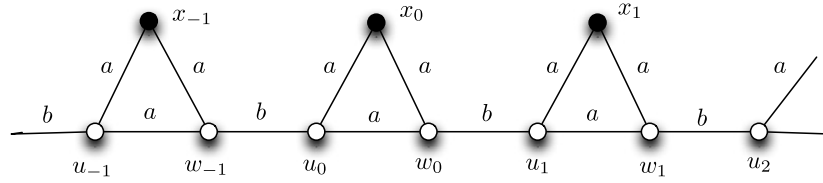


Fig. 5. An infinite model with states  $\{u_i, w_i, x_i \mid i \in \mathbb{Z}\}$  for Theorem 6, item 4. Black dots denote worlds where  $p$  is true.

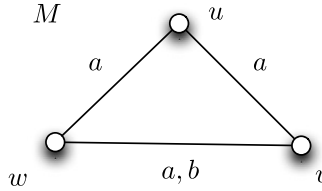


Fig. 6. In this model  $M$  all states satisfy the same atoms.

2. This follows from Example 3 on page 145.
3. Take the model  $M$  of Fig. 6 (recall that we do not draw reflexive arrows in our figures). That is,  $W = \{u, v, w\}$ ,  $R_a = W \times W$ ,  $R_b = \{(v, w), (w, v)\} \cup \{(x, x) \mid x \in W\}$ ,  $I$  interprets  $\succcurlyeq$  as expected and  $V$  is such that  $V(p)$  is either  $W$  or  $\emptyset$ , for all  $p \in \pi$  (so all worlds agree on propositional variables). An easy induction on formulas  $\varphi$  from  $\mathcal{L}_{\text{CEL}}$  shows that

$$M, w \models \varphi \text{ iff } M, u \models \varphi \text{ iff } M, v \models \varphi$$

This is true by definition for  $\varphi \in \pi$ , and it is also true for relational atoms:  $a \succcurlyeq b$  is true in all worlds, and  $b \succcurlyeq a$  is true in no world. The cases for negation, conjunction and knowledge formulas then follow directly. To conclude, we have  $(M, w) \models a \succ b$  but  $\Phi(w)$  is false in  $M$ .

4. See Fig. 5. We will not formally define the model  $M$  depicted in this figure, but it is easy to see that for all  $i, j \in \mathbb{Z}$ , and all  $\varphi$ , we have  $M, x_i \models \varphi$  iff  $M, x_j \models \varphi$ . This is because the two models  $(M, x_i)$  and  $(M, x_j)$  are not just bisimilar, they are isomorphic! From this it follows that for all  $\varphi$ ,  $(M, w_0) \models K_a \varphi \rightarrow K_b \varphi$ . Because if  $(M, w_0) \models K_a \varphi$ , then  $\varphi$  holds in  $w_0, x_0$  and  $u_0$ , so by the claim just mentioned we have that  $\varphi$  holds in  $w_0$  and  $u_1$ , and hence  $(M, w_0) \models K_b \varphi$ .  $\square$

It would be interesting to study under which conditions the counter-examples that perfectly link  $\succcurlyeq$  and ‘knowing at least as much’, or between  $\succ$  and ‘knowing more’ fade away, in other words, under which circumstances there is a perfect match between for instance  $\succ$  and ‘knowing more’. A natural condition for this seems to be the case where, for any agent  $a$ , and any world  $w$ , the set  $R_a(w)$  is definable in  $\mathcal{L}_{\text{CEL}}$ , that is, where there is a formula  $\chi_{a,w}$  such that

$$M, v \models \chi_{a,w} \text{ iff } v \in R_a(w)$$

It is easy to see how this addresses the negative results in Theorem 6. For item 2 for instance, if  $\Psi(w)$  holds, we have in particular  $M, w \models K_a \chi_{a,w} \rightarrow K_b \chi_{a,w}$  and hence if  $R_b wu$ , then  $M, u \models \chi_{a,w}$  and hence  $u \in R_a(w)$ , so that  $M, w \models a \succcurlyeq b$ .

Examples where definability of successors is guaranteed are for instance obtained if we restrict our attention to the case where the language has only *finitely many atoms*, and where the models are both *image finite* (every world has only finitely many successors) and *strongly extensional* (see [3], this notion is also sometimes called *bisimulation contraction*), items 2 to 4 of Theorem 6 would not allow for a counter-example. Intuitively, a model is strongly extensional if it cannot contain fewer worlds without changing its information content, i.e.: removing a state would mean changing the truth of some formula in some other state. We leave a precise characterisation of classes of models where  $\succcurlyeq$  and ‘knowing at least’ exactly match, for future research.

In Proposition 1 we showed that some simple properties of  $\succcurlyeq$  can be already derived in **CL**. An obvious question is whether adding knowledge properties to the logic enables us to derive more properties of  $\succcurlyeq$ , in particular, whether it induces interaction properties between the  $\succcurlyeq$ -operators and the knowledge operators. This is indeed the case.

**Theorem 7.** Let  $a$  and  $b$  agents. Then the following are validities in **CEL**.

1.  $a \succcurlyeq b \leftrightarrow K_b(a \succcurlyeq b)$
2.  $\neg(a \succcurlyeq b) \leftrightarrow K_b \neg(a \succcurlyeq b)$
3.  $a \succ b \leftrightarrow K_b(a \succ b)$
4.  $a \succ b \leftrightarrow (K_b(a \succcurlyeq b) \wedge K_a \neg(b \succcurlyeq a))$
5.  $K_b(a \succcurlyeq b) \vee K_b \neg(a \succcurlyeq b)$

**Proof.** In the first four cases, the right-to-left direction follows from the fact that we have reflexive models (or, equivalently, since we have axiom **T**).

1. Let  $(M, w)$  be arbitrary and assume  $M, w \models a \succ b$ . To prove  $M, w \models K_b(a \succ b)$ . Therefore, let  $v$  be such that  $R_b wv$ , then to prove that  $M, v \models a \succ b$ . Let  $u$  be such that  $R_b vu$ , then to prove that  $R_a vu$ . From  $R_b wv$  and  $R_b vu$  follows  $R_b wu$ , as  $R_b$  is transitive. From assumption  $M, w \models a \succ b$  and  $R_b wv$  follows  $R_a wv$ , and from  $M, w \models a \succ b$  and  $R_b wu$  follows  $R_a wu$ . From  $R_a wv$  and  $R_a wu$  and the Euclidicity of  $R_a$  now follows  $R_a vu$ .
2. This follows from the first item, using modal logical reasoning. Writing  $\alpha$  for  $a \succ b$ , the first item implies that we have  $\models \alpha \rightarrow K_b \alpha$ , and hence  $\models \neg K_b \alpha \rightarrow \neg \alpha$ . Applying necessitation and the  $K$ -axiom to the latter, we obtain  $\models K_b \neg K_b \alpha \rightarrow K_b \neg \alpha$  (\*). Since our models are reflexive, we have  $\models \neg \alpha \rightarrow \neg K_b \alpha$ . Applying negative introspection to the latter yields  $\models \neg \alpha \rightarrow K_b \neg K_b \alpha$ . Using this with (\*) finally gives  $\models \neg \alpha \rightarrow K_b \neg \alpha$ , i.e.,  $\models \neg(a \succ b) \rightarrow K_b \neg(a \succ b)$ .
3. This follows from the previous two items (and hence, since 2 follows from 1, it follows from 1). Suppose  $(M, w) \models a > b$ . This implies  $(M, w) \models a \succ b$  and hence, by item 1, then  $(M, w) \models K_b(a \succ b)$ . Since  $(a > b)$  implies  $\neg(b \succ a)$ , from item 2 we infer that  $M, w \models K_a \neg(b \succ a)$ . Since  $M, w \models a \succ b$  we have  $M, w \models K_a \psi \rightarrow K_b \psi$  for all  $\psi$ , and hence  $M, w \models K_b \neg(b \succ a)$ . In sum, we have  $M, w \models K_b(a \succ b) \wedge K_b \neg(b \succ a)$ , which, by modal reasoning, yields  $M, w \models K_b(a > b)$ .
4. This follows from item 1, item 2, and the definition of  $a > b$ : note that the latter is defined as  $(a \succ b) \wedge \neg(b \succ a)$ , from which, by item 1 we derive  $K_b(a \succ b) \wedge \neg(b \succ a)$ , and, by item 2,  $K_b(a \succ b) \wedge K_a \neg(b \succ a)$ .
5. Again, this follows directly from items 1 and 2, together with the validity of  $(a \succ b) \vee \neg(a \succ b)$ .  $\square$

Theorem 7 states some facts about  $\succ$  on S5-models that at first sight might seem remarkable. Roughly speaking, they imply that if an agent  $b$  has at least the same (or more) information than agent  $a$ , agent  $b$  will know this. In particular, item 1 of the theorem says that if  $b$  considers at most the states possible that  $a$  considers possible,  $b$  knows this. Item 3 states that if  $b$  only considers a strict subset of  $a$ 's alternatives,  $b$  knows this as well! This is particularly striking in cases where  $a$  and  $b$  in fact know the same (this is for instance true in  $(M, w)$  of Fig. 6). In this model, note that agents  $a$  and  $b$  know the same in  $(M, w)$ . However, although both agents know exactly the same, both agents also know that agent  $b$  considers less states possible than agent  $a$ ! Item 2 then says that if  $b$  does *not* consider at most the states possible that  $a$  considers possible,  $b$  must know this. In sum, the first three items seem to suggest that no matter whether (i)  $b$  has at least the information as  $a$  has, or (ii) has more information than  $a$ , or (iii) does not have at least the information that  $a$  has,  $b$  will know this! Item 5 states that, for any agent  $b$ , if he compares his information state with an arbitrary agent  $a$ , then  $b$  will know whether he has at least the same information as  $a$  or not. The following representation of S5-models and the alternative notation when dealing with equivalence relations may also help the reader.

**Remark 2.** We give an alternative proof of Theorem 7 items 1 and 3, stressing the fact that we are dealing with equivalence relations. For an S5-model  $M = \langle W, R, V, I \rangle$ , a world  $w \in W$ , and an agent  $a$ , let us write  $[a]_w$  for  $\{v \in W \mid R_a wv\}$ . Note that, when  $v \in [w]_a$ , then  $[w]_a = [v]_a$ . In the model  $M$  of Fig. 7, we have depicted four states,  $w, v, x$  and  $z$ . The set  $W$  is partitioned in  $[w]_a$  and  $[z]_a$  (note that in fact there does not need to exist a  $z \notin [w]_a$ , in which case we would only have one equivalence class for  $a$ ). For agent  $b$ , we have only depicted one equivalence class  $[w]_b$ , it does not matter how the other classes for  $b$  are splitting up  $W$ . The existence of world  $x$  is only guaranteed in the second item below.

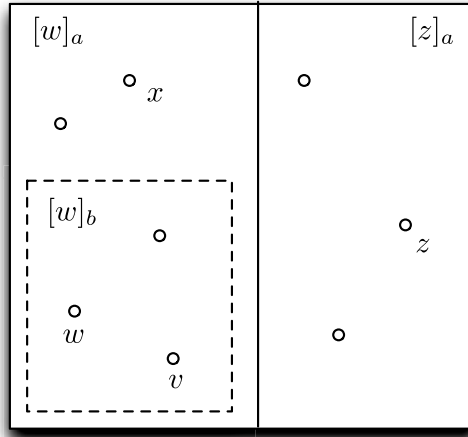
1. Let  $M, w \models a \succ b$ . This means that  $[w]_a \supsetneq [w]_b$  (1). In order to prove  $M, w \models K_b(a \succ b)$  we need to show that for any  $v$  such that  $v \in [w]_b$ , we have that  $M, v \models a \succ b$ , that is,  $[v]_a \supsetneq [v]_b$ . Since  $v \in [w]_b$ , we have  $[w]_b = [v]_b$ , which with (1) gives  $[w]_a \supsetneq [v]_b$  (2). Since  $v \in [w]_b$ , with (1) we derive that  $v \in [w]_a$ , and hence  $[v]_a = [w]_a$ . With (2) this gives  $[v]_a \supsetneq [v]_b$ .

The proof of item 3 of Theorem 7 is almost identical to that of the previous item: replace each occurrence of  $\succ$  by  $>$  and each occurrence of  $\supsetneq$  by  $\supset$ .

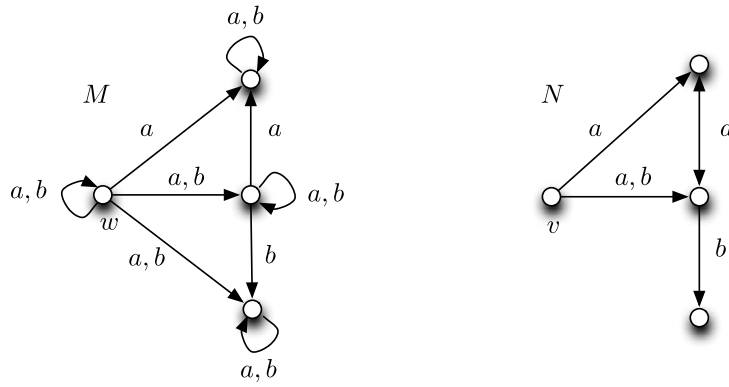
Where our first proof of Theorem 7 shows that transitivity and Euclidicity are sufficient to demonstrate the  $\rightarrow$ -direction of the items 1 and 3, we finally show that they are also necessary: see Fig. 8. In model  $M$ , we have  $(M, w) \models M, w \models a > b \wedge \neg K_b(a \succ b)$  which, shows that both items 1 and 3 of Theorem 7 fail on reflexive transitive models. In  $N$ , we have  $(N, v) \models a > b \wedge \neg K_b(a \succ b)$ , showing that Euclidicity on its own cannot guarantee Theorem 7 to hold. (There can be no reflexive Euclidean model refuting Theorem 7, since those two conditions together ensure that the accessibility relation is also transitive, i.e., an equivalence.)

Concentrating now on the agent  $a$  that has less information (i.e., who considers at least the states possible that  $b$  considers possible), it turns out that  $a$  does not need to be aware of this. It seems somewhat 'unfair' that if  $a$  knows *less* than  $b$ ,  $a$  does not necessarily know that – at least in a (fair!) game-like setting one would expect the opposite: player  $b$  may have an advantage (more information, more knowledge), but in order to place his bets player  $a$  should at least be aware of his disadvantage (ignorance) compared to  $b$ . On the other hand, considering that these are typically incomplete information games, from  $a$ 's perspective it may be gambling the possibility of being less informed than  $b$  against the possibility of knowing more, which sounds more 'fair'.

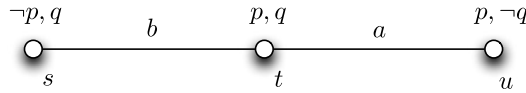
**Example 4.** Consider the model of Fig. 9 consisting of three states, where  $a$  cannot distinguish  $t$  from  $u$  whereas  $b$  cannot distinguish  $s$  from  $t$ . In fact  $u$  is the case. In  $u$  it is true that  $a \succ b$  – in fact  $a$  knows indeed less than  $b$ , but  $a$  considers



**Fig. 7.** In this model  $M$  we have depicted two equivalence classes for  $a$ , but only one for agent  $b$ .



**Fig. 8.** In both  $(M, w)$  and  $(N, v)$ , formula  $a \succ b \wedge \neg K_b(a \succ b)$  holds.



**Fig. 9.** An S5-model with three states. The valuation is given above each state.

it possible that  $\neg(a \succ b)$ , in which case the advantage of  $b$  would be less. Possibly  $a$  is willing to bet for the latter, even though the former is really the case. This is like playing bridge and noting the signals exchanged between the opponents, that if true may have revealed their better hand of cards, but they may also have been explicitly misinforming you by their signals in order to confuse you. In any case, our model has  $(M, u) \models a \succ b \wedge \neg K_a(a \succ b)$ : agent  $a$  is at most as informed as  $b$  without knowing it. Note that we even have  $(M, u) \models a \succ b \wedge \neg K_a(a \succ b)$ : agent  $a$  is less informed than  $b$  without knowing this.

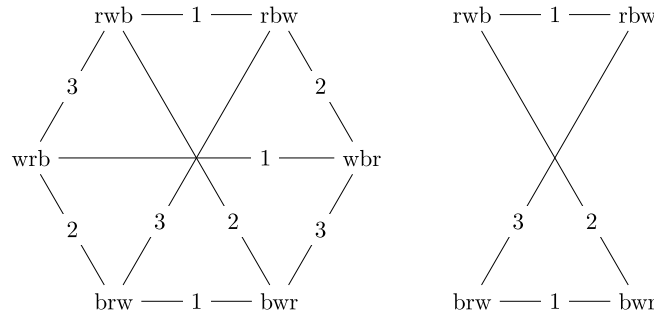
It is of course well possible that each agent has a secret. In Fig. 9 for instance, we have

$$M, t \models (K_a p \wedge \neg K_b p) \wedge (K_b q \wedge \neg K_a q)$$

Before looking at an example in **CEL**, consider the following property:  $\neg K_a \varphi \wedge \bigwedge_{i \in A} ((a \succ i) \rightarrow K_i \varphi)$ . This expresses that  $a$  does not know  $\varphi$ , but anybody who would know even a little bit more would know  $\varphi$ . In state  $u$  of the example above we have that  $a$  does not know that  $q$  is false, but  $b$ , who knows ‘a little bit’ more (who considers only the actual state as possible, one less than the two states considered possible by  $a$ ) knows that  $q$  is false. So we have, in  $u$ :

$$\neg K_a \neg q \wedge (a \succ b \rightarrow K_b \neg q)$$

This notion of ‘knowing even a little bit more’ seems of independent interest on infinite models, where the difference between  $a$ ’s ignorance and any other agent’s knowledge may be like the difference between a closed set and any open approximation of that set.



**Fig. 10.** Two epistemic models  $M$  and  $N$ : in  $(N, rwb)$  agent 3 knows more than agent 1 and more than agent 2.

#### 4.3. Comparative epistemic logic with public announcements

Being able to locally express that agent  $a$  knows at least what agent  $b$  knows is especially valuable if one can somehow reason about *change* of knowledge, allowing one to express that it comes about that one player gets to know at least what another knows. Dynamic epistemic logic ([27]) is a powerful formalism to reason about change within the object language. Without going into the details of the logic, we will here focus on a particular case of dynamic epistemic logic, namely public announcements, thereby focus on the semantical aspect of model restriction.

First, consider the two epistemic models in Fig. 10. Call the models  $M$  (left) and  $N$  (right). Model  $M$  represents a situation where three players 1, 2 and 3 each hold a card from a deck of cards  $\text{Cards} = \{r, w, b\}$ . A state with a name  $xyz$  represents a situation where player 1 holds card  $x$ , player 2 holds  $y$  and player 3 holds  $z$ . We also assume atoms  $r_i, w_i, b_i$  ( $i \leq 3$ ) where for instance  $r_2$  is true in those situations where 2 holds the red card. Players only see their own card, and not the card of the others (and this is common knowledge). Note that both models are image finite and strongly extensional: the worlds in each model already differ in the valuations. This implies that we can here identify  $\succsim$  with ‘knowing at least as much as’, and  $\succ$  with ‘knowing more’.

In  $M$  we have for instance that

$$M, rwb \models r_1 \wedge K_1 r_1 \wedge \neg K_2 r_1 \wedge K_1 \neg K_2 r_1$$

i.e., if the deal is  $rwb$ , then 1 holds the red card, he knows this, but 2 does not know it, and, finally, 1 knows that 2 does not know that 1 holds the red card.

Since every player knows a fact that the others don’t know (i.e., the face of their own card) we also have

$$M \models \neg(1 \succsim 2) \wedge \neg(2 \succsim 1) \wedge K_1 \neg(1 \succ 2) \wedge K_1 K_2 \neg(1 \succ 2)$$

which says that the information of players 1 and 2 is incomparable, and 1 knows that 2 does not know more than him: 1 even knows that 2 knows that 2 does not know more than 1.

However, now consider model  $N$ , which can be obtained from  $M$  as the effect of a public announcement that player 1 does not hold the white card, i.e.,  $\neg w_1$ . In other words,  $N$  is obtained from  $M$  by leaving out all those states where the announcement (that  $\neg w_1$ ) is false. Said differently,  $N$  is the result of  $M$  when leaving out the states  $wrb$  and  $wbr$ . The notation for this is  $N = M | \neg w_1$  ( $N$  is  $M$  restricted to the  $\neg w_1$  worlds). We then obtain

$$N, rwb \models 1 \succ 3 \wedge 2 \succ 3 \wedge M_1(3 \succ 2)$$

This describes the situation that when the card deal is  $rwb$ , and somebody announces in public that 1 does not hold the white card, that after that announcement, 3 knows more than 1 and than 2 (they both knew already the fact that is announced), and 1 considers it possible that 2 holds the white card, in which situation  $3 \succ 2$  would hold after the announcement.

Let us now consider a little closer how a system of comparative epistemic logic combined with public announcement logic would look like. First of all, this would involve a language, which we call  $\mathcal{L}_{\text{CEL}+\text{PAL}}$  defined as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_a \varphi \mid a \succsim b \mid [\varphi]\varphi$$

where  $[\varphi]\psi$  is read as: ‘after the public announcement of  $\varphi$ ,  $\psi$  is the case’. We have already discussed the semantics for this: if  $M = \langle W, R, V, I \rangle$ , then

$$(M, w) \models [\varphi]\psi \quad \text{iff} \quad (M, w) \models \varphi \text{ implies } (M|\varphi, w) \models \psi$$

where  $M|\varphi$  is the model  $M$  restricted to those worlds that verify  $\varphi$ .

In such a combination, one would enable to communicate what in standard public announcement logic would require an infinite amount of communication (namely, as above, public announcement of  $K_a \varphi \rightarrow K_b \varphi$  for all formulas  $\varphi$  in the language!). The public announcement  $a \succsim b$  has the effect that the local property that  $b$  knows at least what  $a$  knows

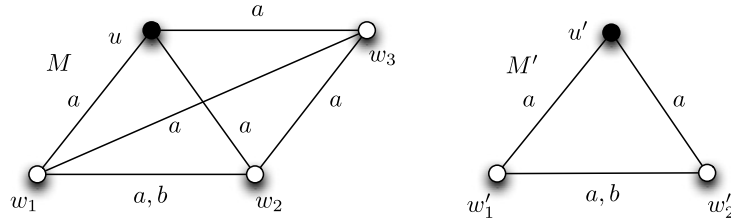


Fig. 11. Two S5-models. States where  $p$  holds are black.

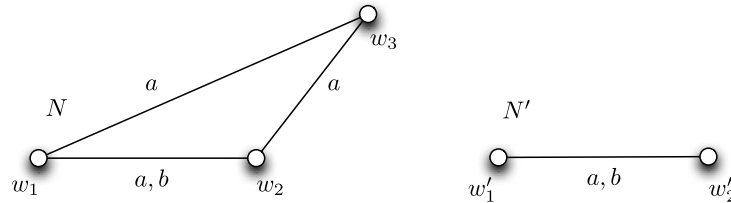


Fig. 12. Two S5-models, the result of announcing  $\neg p$  in the models of Fig. 11.

becomes a global property. It is easy to see that  $[a \succcurlyeq b](K_a \varphi \rightarrow K_b \varphi)$  is a validity, or, in other words, after an announcement with  $a \succcurlyeq b$ ,  $R_b \subseteq R_a$  is common knowledge. (If  $[a \succcurlyeq b](K_a \varphi \rightarrow K_b \varphi)$  is a validity then also, as usual,  $[a \succcurlyeq b]C_A(K_a \varphi \rightarrow K_b \varphi)$ .) The semantics for such a public announcement would be a standard model restriction to the worlds satisfying the announced formula, in this case, to the worlds satisfying  $a \succcurlyeq b$ :  $M|a \succcurlyeq b = \{s \in M \mid M, s \models a \succcurlyeq b\}$ .

The axiomatisation of public announcements in S5 is obtained by so-called *reduction axioms* (see [27] for a discussion). Those axioms explain how public announcement can be reduced to statements in S5 that do not involve such announcements. For instance, the reduction axiom for knowledge says that  $[\varphi]K_i \psi$  ('after announcement of  $\varphi$ , agent  $i$  knows  $\psi$ ') is equivalent to  $\varphi \rightarrow K_i[\varphi]\psi$ . Using this axiom together with the other axioms (for conjunction and negation), we can 'push' the announcements operators  $[\cdot]$  'inside' the subformulas, and eventually, using the axioms for announcement and atoms, get rid of them. This shows that S5 plus public announcements is equivalent to S5 itself.

So if we would be able to come up with a reduction principle for announcements and  $\succcurlyeq$ , we would have a complete logic for a language with knowledge, announcements and the comparison operator  $\succcurlyeq$ . However, there is no reduction axiom for  $[\varphi]a \succcurlyeq b$ . Even stronger, we will now show that by adding public announcements to **CEL** we obtain a system that is more expressive than **CEL** itself.

**Theorem 8.** *CEL with public announcements is more expressive than CEL.*

**Proof.** We will employ a results about coalitional bisimulations, Theorem 4 (first item). Consider the two models  $M = \langle W, R, V, I \rangle$  and  $M' = \langle W', R', V', I' \rangle$  depicted in Fig. 11: the atoms  $p$  is true at states  $u$  and  $u'$  only, and the accessibility relations for the agents  $a$  and  $b$  can be read off from the figure (note that this is an S5-model: reflexive arrows are not drawn). Formally:  $W = \{u, w_1, w_2, w_3\}$  with  $R_a = W \times W$  and  $R_b$  is the reflexive closure in  $W$  of  $\{(w_1, w_2), (w_2, w_1)\}$ , and  $V(p) = \{u\}$ . Similarly,  $W' = \{u', w'_1, w'_2\}$  with  $R'_a = W' \times W'$ ,  $R'_b$  is the reflexive closure on  $W'$  of  $\{(w'_1, w'_2)\}$  and  $V'(p) = \{u'\}$ . Both interpretations  $I$  and  $I'$  link  $\succcurlyeq$  with the superset-relation. We claim the following:

$$\forall \varphi \in \mathcal{L}_{\mathbf{CEL}} (M, w_1) \models \varphi \quad \text{iff} \quad (M', w'_1) \models \varphi \quad (9)$$

This is proven by showing that there exists a (natural) extended bisimulation between  $M, w$  and  $M', w'_1$ , which is

$$\mathcal{R} = \{(u, u')\} \cup \{(w_i, w'_j) \mid i \leq 3, j \leq 2\}$$

To verify that  $\mathcal{R}$  respects propositional variables is straightforward:  $p$  is true in  $u$  and  $u'$ , and false in all other worlds. For relational atoms,  $b \leq a$  is true in all seven worlds, and  $a \leq b$  is false everywhere. (Moreover,  $b \leq a$  is interpreted in  $x$  as  $R_b(x) \subseteq R_a(x)$ , which makes the extended bisimulation natural.) The 'forth' and 'back' conditions of  $\mathcal{R}$  are also straightforward to check. (9) now follows with Lemma 6.

Now, consider the models  $N$  and  $N'$  of Fig. 12. The situation  $(N, w_1)$  is the result of publicly announcing  $\neg p$  in  $(M, w_1)$ , whereas  $(N', w'_1)$  is the result of announcing  $\neg p$  in  $(M', w')$ . Yet, we have  $(N, w_1) \models a > b$  but  $(N', w'_1) \models \neg(a > b)$ . This implies

$$(M, w_1) \models [\neg p](a > b) \quad \text{while} \quad (M', w'_1) \not\models [\neg p](a > b)$$

In other words, although for the language  $\mathcal{L}_{\mathbf{CEL}}$  the models  $(M, w_1)$  and  $(M', w'_1)$  are the same, in the language  $\mathcal{L}_{\mathbf{CEL}+\mathbf{PAL}}$ , we find a formula that distinguishes  $(M, w_1)$  and  $(M', w'_1)$ .  $\square$

So Theorem 8 in effect shows that the elimination process that underlies the semantics of a public announcement, can change the  $\succsim$  properties between ‘similar’ models in such a way that they lose their similarity. In particular, where the presence of world  $w_3$  in model  $M$  is irrelevant for what is true in  $w_1$  (this is what (9) says), because it ‘survives’ the announcement  $\neg p$ , the presence of  $w_3$  in the resulting model makes that  $w_1$  has more  $a$ -successors than  $b$ -successors, which would not have been so (in  $N'$ ) if  $w_3$  was left out in the first place (like in  $M'$ ).

#### 4.3.1. Adding coalitional epistemic operators

The notion of ‘knowing at least as’ for individuals has at least two extensions to that of groups. Let  $C$  and  $D$  to be two coalitions. We can interpret  $C \succsim^\cap D$  in  $w$  as  $\bigcap_{c \in C} R_c(w) \supseteq \bigcap_{d \in D} R_d(w)$  and  $C \succsim^\cup D$  in  $w$  as  $\bigcup_{c \in C} R_c(w) \supseteq \bigcup_{d \in D} R_d(w)$ . Then,  $C \succsim^\cap D$  would mean: ‘the distributed knowledge of  $D$  is at least that of  $C$  (in  $w$ )’, and  $C \succsim^\cup D$  would mean ‘what everybody in  $D$  knows is at least what everybody in  $C$  knows (in  $w$ )’. For instance, the sentence ‘Steve knows at least what his parents know’ would have the following three interpretations:  $(p_1 \succsim s) \wedge (p_2 \succsim s)$  (‘Steve knows at least what each of his parents knows’) and  $\{p_1, p_2\} \succsim^\cup \{s\}$  (‘Steve knows at least what both of his parents know’) and  $\{p_1, p_2\} \succsim^\cap \{s\}$  (‘Steve knows at least what his parents distributively know’).

For notions of group knowledge, many other options present themselves. It is well known that common knowledge of a coalition  $D$ , written  $C_D$ , semantically corresponds to the transitive closure  $R_D^*$  of the union of the individual relations [7]  $R_i$  ( $i \in D$ ). So one could add primitives like  $D^* \succsim F^*$  indicating that the common knowledge of coalition  $D$  is a subset of the common knowledge of group  $F$ . And the notion of group knowledge on both sides of  $\succsim$  do not have to coincide either:  $D^* \succsim F^E$  for instance might read: ‘currently, all what is common knowledge in coalition  $D$ , is known by everybody in  $F$ ’.

#### 4.3.2. Only knowing

There is a rich literature on ‘only knowing’ also in the multi-agent context see for instance [13], the overview paper [26] or the recent [2]. Although related to the issues that **CEL** addresses, there are also differences: in only knowing, one tries to characterise the minimal amount of knowledge of an agent, given he knows a certain fact  $\varphi$ . In **CEL**, the emphasis is on comparing one agent’s knowledge to another agent’s.

## 5. Conclusion

We have presented a flexible way to deal locally with quantification over formulas. In particular, we have shown how, under some mild conditions, in a modal logic that extends **K** with some canonical axioms, one can add a number of relational atoms, for each of them an axiom and an inference rule, such that the logic is complete for the class of models that interpret the atom as a first order property of the underlying frame. We argued that this presents many opportunities to express properties concerning the knowledge or beliefs of agents in a local way, so that they are only true now, or as a belief or knowledge of some specific agents.

In more detail we investigated the case of ‘comparative epistemic logic’, for the relational atom  $\succsim$  such that  $a \succsim b$  informally stands for ‘ $b$  knows at least as much as  $a$ ’. Although we focussed on epistemic and doxastic logics, our technique is applicable in dynamic settings as well. On our agenda is to study how our framework behaves in a dynamic epistemic logic setting.

## Acknowledgements

We thank the three reviewers of *Artificial Intelligence* for their very helpful comments and suggestions. We also thank Frank Wolter for providing useful comments on a pre-final version of the paper. Hans van Ditmarsch is also affiliated to IMSc, Chennai, India, as research associate.

## References

- [1] A. Baltag, L.S. Moss, S. Solecki, The logic of common knowledge, public announcements, and private suspicions, in: I. Gilboa (Ed.), TARK 98, 1998, pp. 43–56.
- [2] Vaishak Belle, Gerhard Lakemeyer, Multi-agent only-knowing revisited, CoRR, abs/1009.2041, 2010.
- [3] P. Blackburn, M. de Rijke, Y. Venema, Modal Logic, vol. 53, Cambridge University Press, Cambridge, 2001.
- [4] P. Blackburn, J. van Benthem, F. Wolter (Eds.), Handbook of Modal Logic, Elsevier, Amsterdam, The Netherlands, 2007.
- [5] P. Cohen, H. Levesque, Intention is choice with commitment, Artificial Intelligence 42 (1990) 213–261.
- [6] M. de Rijke, Extending modal logic, PhD thesis, ILLC, University of Amsterdam, 1993.
- [7] R. Fagin, J. Halpern, Y. Moses, M. Vardi, Reasoning About Knowledge, MIT Press, 1995.
- [8] D. Gabbay, An irreflexivity lemma with applications to axiomatisations of conditions on tense frames, in: U. Monnich (Ed.), Aspects of Philosophical Logic, Reidel, Dordrecht, 1981, pp. 67–89.
- [9] R. Goldblatt, Axiomatising the Logic of Computer Programming, Lecture Notes in Computer Science, vol. 130, Springer, Berlin, 1982.
- [10] R. Goldblatt, Logics of Time and Computation, 2nd ed., CSLI Lecture Notes, vol. 7, CSLI Publications, Stanford, CA, 1987 (revised and expanded 1992).
- [11] R. Goldblatt, Mathematics of Modality, CSLI Lecture Notes, vol. 43, CSLI Publications, Stanford, CA, 1993.
- [12] V. Goranko, Axiomatizations with context rules of inference in modal logic, Studia Logica (1998) 179–197.
- [13] J. Halpern, G. Lakemeyer, Multi-agent only knowing, Journal of Logic and Computation (1996) 251–265.
- [14] J.Y. Halpern, Y. Moses, Knowledge and common knowledge in a distributed environment, JACM 37 (3) (1990) 549–587.



- [15] D. Harel, Dynamic logic, in: D. Gabbay, F. Guenther (Eds.), *Handbook of Philosophical Logic. Vol. II – Extensions of Classical Logic*, in: Synthese Library, vol. 164, D. Reidel Publishing Company, Dordrecht, The Netherlands, 1984, pp. 497–604.
- [16] J. Hintikka, *Knowledge and Belief. An Introduction into the Logic of the Two Notions*, Cornell University Press, Ithaca, NY, 1962.
- [17] A. Lomuscio, H. Qu, M. Sergot, M. Solanki, Verifying temporal and epistemic properties of web service compositions, in: *Proceedings ICSOC07*, in: LNCS, vol. 4749, 2007, pp. 456–461.
- [18] J.-J.Ch. Meyer, W. van der Hoek, *Epistemic Logic for AI and Computer Science*, Cambridge University Press, 1995.
- [19] R.C. Moore, Reasoning about knowledge and action, in: *IJCAI-77*, Morgan Kaufmann, Cambridge, MA, 1977.
- [20] A.S. Rao, M.P. Georgeff, Modeling rational agents within a BDI-architecture, in: R. Fikes, E. Sandewall (Eds.), *KR&R-91*, Morgan Kaufmann, 1991, pp. 473–484.
- [21] K. Segerberg, A model existence program in infinitary propositional modal logic, *Journal of Philosophical Logic* 23 (4) (1994) 337–367.
- [22] B. ten Cate, *Model theory for extended modal languages*, PhD thesis, University of Amsterdam, 2005.
- [23] J. van Benthem, *Modal correspondence theory*, PhD thesis, University of Amsterdam, 1976.
- [24] J. van Benthem, *Modal Logic and Classical Logic*, Bibliopolis, Napoli, 1985.
- [25] W. van der Hoek, M. Pauly, Modal logic for games and information, in: P. Blackburn, J. van Benthem, F. Wolter (Eds.), *Handbook of Modal Logic*, Elsevier, Amsterdam, The Netherlands, 2007, pp. 1077–1148.
- [26] W. van der Hoek, E. Thijsse, A general approach to multi-agent minimal knowledge: With tools and samples, *Studia Logica* 72 (1) (2002) 61–84.
- [27] H. van Ditmarsch, W. van der Hoek, B. Kooi, *Dynamic Epistemic Logic*, Springer, Berlin, 2007.
- [28] H. van Ditmarsch, W. van der Hoek, B. Kooi, Knowing more – towards a local correspondence theory, in: C. Boutilier (Ed.), *Proceedings IJCAI-09*, 2009, pp. 955–960.
- [29] H. van Ditmarsch, W. van der Hoek, B. Kooi, Reasoning about local properties in modal logic, in: K. Tumer, P. Yolum, L. Sonenberg, P. Stone (Eds.), *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, 2011, pp. 711–718.
- [30] Y. Venema, Derivation rules as anti-axioms in modal logic, *Journal of Symbolic Logic* (1993) 1003–1034.
- [31] M. Wooldridge, *An Introduction to Multiagent Systems*, John Wiley & Sons, 2002.